DECEMBER
1 9 5 6
VOLUME 20
NUMBER 6

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

Published by the American Psychological Association, Inc.
at Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth St. N.W., Washington, D.C.

Edited by Laurance F. Shaffer

Teachers College, Columbia University

Advisory Editors: Donald W. Fiske, *University of Chicago.* Ann Magaret Garner, *University of Illinois.* Harrison G. Gough, *University of California.* William A. Hunt, *Northwestern University.* E. Lowell Kelly, *University of Michigan.* Morris Krugman, *Board of Education, New York, N. Y.* Bertha M. Luckey, *Board of Education, Cleveland, Ohio.* Fred McKinney, *University of Missouri.* Anne Roe, *New York, N. Y.* Carl R. Rogers, *University of Chicago.* Nevitt Sanford, *University of California.* Edward Joseph Shoben, Jr., *Teachers College, Columbia University.* George J. Wischner, *University of Pittsburgh.* J. R. Wittenborn, *Rutgers University.*

Lorraine Bouthilet, *Managing Editor*

The *Journal of Consulting Psychology* is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

Editorial Office

Manuscripts submitted for publication and correspondence on editorial matters should be sent to the Editor, Laurance F. Shaffer, 525 West 120th St., New York 27, N. Y. All manuscripts should be submitted *in duplicate* to facilitate referral to the Advisory Editors. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

Style of manuscripts. Manuscripts should be prepared in the style described in the Publication Manual of the American Psychological Association (*Psychological Bulletin*, 1952, 49, 389–449). The Manual may be purchased from the Washington office of the Association for $1.00. Otherwise acceptable manuscripts which depart from the standards of the Manual are returned for revision and retyping.

Brief Reports. A one-page condensation of an acceptable research report may be given early publication without charge if the author agrees (*a*) to make 100 mimeographed copies of the full report available to those who request it, (*b*) to send a copy of the full report to the Editor for deposit in the American Documentation Institute, and (*c*) not to have the full report published elsewhere. For detailed instructions on how to prepare a Brief Report, see the back of the *Contents* page of this Journal or write to the Editor.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. The charge to the author for such early publication is $20.00 per printed page.

Reprints. Authors receive fifty reprints gratis of all articles except early publication articles and Brief Reports. Additional quantities may be ordered when returning proof.

Business Office

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters, should be sent to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C. Address changes must reach the subscription office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

The *Journal of Consulting Psychology* is published bimonthly in February, April, June, August, October, and December, the yearly volume comprising approximately 480 pages. The subscription per year is $8.00, foreign $8.50, single number $1.50.

Entered as second-class matter, March 21, 1955, at the post office at Lancaster, Pa., under the act of March 3, 1879. Acceptance for mailing at special rate of postage provided for in Section 538, Act of February 28, 1925, authorized May 17, 1940.

Copyright © 1956 by the American Psychological Association, Inc.

JOURNAL OF CONSULTING PSYCHOLOGY

December, 1956 Vol. 20, No. 6

Contents

Second-Order Personality Factors in the Questionnaire Realm: *Raymond B. Cattell* - - 411

Factors, Response Bias, and the MMPI: *Eugene H. Barnes* - - - - - - - - - - - 419

The Personality of Judges and their Agreement with Experts in Judging Clinical Case Histories: *A. W. Bendig* - 422

Comparison of Two Methods of Electroshock in Their Effect on Cognitive Functions: *Barbara S. Kendall, Warren B. Mills, and Thomas Thale* - - - - - - - - - - 423

Group Psychotherapy with Chronic Psychotic Soiling Patients: *John E. Tucker* - - - - 430

The Short-Term Effects of Chlorpromazine on Psychiatric Patients: *William D. Winter and Wilbur K. Frederickson* - 431

Use of the Proverbs Test for Differentiating Schizophrenics from Normals: *Donald R. Gorham* - 435

The Ability of Schizophrenics to Modify Responses in an Interpersonal Situation: *M. David Diamond* - 441

Concept Learning in the Educationally Retarded Child of Normal Intelligence: *Nissim M. Levy and Joseph M. Cuddy* - 445

The TAT and Antisocial Behavior: *Kenneth Purcell* - - - - - - - - - - - - - - 449

Situational Influences on Rorschach Responses: *Edith M. Henry and Julian B. Rotter* - - 457

Affective Control and Rorschach Human Movement Responses: *Julian Meltzoff and Dorothy Litwin* - 463

The Use of the Kaleidoscope as an Adjunct to the Rorschach: *Emanuel Starer* - - - - 466

The Use of Human Figure Drawings in the Diagnosis of Organic Pathology: *Marvin Reznikoff and Donald Tomblen* - 467

Certain Formal Characteristics of Human Figure Drawings by Institutionalized Aged and by Normal Children: *Martin Lakin* - - - - - - - - - - - - - - - - - 471

Cue-Counting: A Measure of Anxiety in Interviews: *Andrew S. Dibner* - - - - - - - 475

Test Anxiety and Intelligence Test Performance: *I. Zweibelson* - - - - - - - - - - 479

The Relation of the Children's Manifest Anxiety Scale to the Concept of Anxiety as Used in the Clinic: *Robert D. Wirt and William E. Broen, Jr.* - - - - - - - - - - 482

The "Alienation Syndrome" and Adler's Concept of "Distance": *H. L. Ansbacher* - - - 483

Some Comments on Ansbacher's Note on the "Alienation Syndrome" and Adler's Concept of "Distance": *Anthony Davids* - 485

Psychological Test Reviews - 487

Volume Title Page and Contents Appear at the End of this Issue

Brief Reports

The *Journal of Consulting Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.

2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the American Documentation Institute to insure indefinite availability.

3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.

4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, including all matter except the title and the author's lines, must not exceed 70 lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style (1). Headings, tables, and references are avoided or, if essential, must be counted in the 70 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 70-line quota: [1]

[1] An extended report of this study may be obtained without charge from John Doe, 300 Market St., Prospect 6, Mass. (giving the author's full name and address), or for a fee from the American Documentation Institute. Order Document No. —— from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $—— for microfilm or $—— for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Extended report. The full report is prepared in the style specified by the *Publication Manual* (1), except that it may be typed with single spacing for economy in photoduplication by the ADI.

Reference

1. American Psychological Association. Council of Editors. Publication manual of the American Psychological Association. *Psychol. Bull.*, 1952, 49, 389–449.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Second-Order Personality Factors in the Questionnaire Realm

Raymond B. Cattell

Laboratory of Personality Assessment and Group Behavior, University of Illinois

Although there has been much emphasis on the potential importance of broader, second-order personality factors, by Thurstone and others, and even some claims to recognizing particular second-order factors by Eysenck (20) and the present writer (2, 4), there has not existed until this year any very firm basis of evidence for conclusions. The present writer's earlier attempt (4) was admittedly premature, while Eysenck's extensive arguments for a second-order neuroticism factor and "extraversion" factor (20) seem partly based on the misapprehension, mentioned below, that second-order factors can be reliably obtained from factoring correlations among the actual scores on the existing measures for primary factors. There are, indeed, three rather widespread misapprehensions, dealt with below, which may confuse research on second-order factors.

A second-order factor is a factor among oblique primary factors. That is to say, it is obtained by factoring a matrix of correlations among pure measures of primary factors and rotating to simple structure. The second-order factors will be less numerous than first orders, or "primaries," and will tend to cover wider areas of behavior. For example, the general ability factor, on which the IQ is based, covers more than the individual primary abilities of Thurstone, from which it is derived as a second-order factor (30).

The theoretical and practical value of second-order factors, relative to first-order factors, is an issue with partisans on both sides. In the theoretical realm some, like Hsü (25) see second-order factors as "super factors," more permanent and more deeply rooted in physiological or sociological fact than first-orders. Others (6, 24, 33) consider the primaries as the psychologically-real concepts and second-orders only as more transient organizing principles among them. The former is illustrated by the theory that *g* is a function of the total effective number of cortical neurones, but that the primary abilities merely represent particular training investments of this general relation-perceiving capacity, developed in particular sensory or symbolic areas. The latter, i.e., the alternative interpretation, is illustrated by the present writer's earlier explanation that the observed organization of certain personality factors (2) represents only the population correlation produced among them by selection according to the social status conditions of the particular country concerned. (See also Fisher, 21.)

Regardless of theory, the practical psychometrist can predict test performance either from first- or second-order factors. But here, too, there are preferences, as shown, for example, by the British predilection for working with Spearman's *g*—the second-order factor in abilities (33)—and the American preference for primary abilities. If we rule out specifics as too numerous to be useful, prediction from primaries accounts for the greatest possible actual test variance, but from second-order factors there is a wider field of prediction, in terms of one factor (6, 30).

Second-Order Factoring

For the psychologist, as contrasted with the statistician, the importance of factors lies beyond immediate prediction. It concerns isola-

tion of psychological concepts and the development of psychological laws and insights, which extend the understanding of behavior beyond the immediate, static, measurement situation. Here, both the first- and the second-order structure are parts of the total necessary knowledge, as will be exemplified in the present study. However, as indicated in the introductory remarks, the second-order factors must for this purpose meet certain conditions not sufficiently recognized in some current studies.

In the first place, it does not suffice to take any existing test or battery measuring the factors, to intercorrelate the factor scores, and analyze. Today, and for many years to come, the individual factor measures will be substantially impure. Each will contain, in addition to the inevitable specific, some contamination with other factors present in the battery. The primary measures may, also, include in common some factors, e.g., response sets, which have nothing to do with the intended primary measures per se. Consequently the factoring of these actual primary factor test scores, presuming that each is split in two, or increased in number in some way to give a sufficiency of variables for determining the number of common factors involved, will yield the primaries again, together, perhaps, with the spurious common factors among them due to dimensions extraneous to the primary dimensions themselves, e.g., to the response format of the battery.

This theoretical expectation has actually been verified in Thurstone's (31) and Lovell's (27) work on Guilford's factors (24), where, if the battery for each factor is split, and each factor reliability is put in the matrix, the analysis is shown to yield the primaries again with tolerable precision. This result will be readily appreciated by anyone who has seen the complete discrepancies between the correlation between the raw scores of any two factor measures and the correlation found from measuring the cosine of the angle between the factors on their simple-structure plot. The former, dependent on a cluster of items, is pulled all over the place by the addition of fresh items, with their particular contaminations; but the latter remains firm. For the hyperplane is fixed by variables in *many* dimensions that are *not* loaded in the factor in question. However, it must be granted that where factors are already very purely measured by existing tests, with no systematic bias and with effective suppression of peer factors in each factor score, the factoring of actual factor scores, i.e., of one variable for each factor, may yield a reasonable approximation to second-order factors. This seems to have happened with the 16 P.F. test in the large Air Force study (8).

However, the matrix with which a second-order factoring properly begins is the cosine matrix among the simple-structure primary factors. Few oblique rotations to simple structure undertaken in the usual course of determining the form of the primaries are accurate enough for a second-order factoring. Indeed, as French's (22) and other surveys show, perhaps a half of all published factor analyses falter and fall short of their goal—even the goal of defining the first-order factors—for lack of thorough search for simple structure. Even the last couple of over-all, "polishing" rotations, in the usual series of ten to twenty over-all rotations from the original position, will substantially alter the number and nature of second-order factors that would be obtained.

One needs also to be warned against absent-mindedly factoring—as the writer has done before (4)—the matrix of cosines thus achieved, for these are the correlations among the reference vectors. Such a matrix must be inverted to give correlations among the factors. Thus when, as in the usual notation, λ is the obtained transformation matrix, proceeding from the extracted factor matrix, V_o, to the best simple structure matrix, V_s, the required square matrix for second-order factoring is

$$D(\lambda'\lambda)^{-1}D$$

where D is a diagonal matrix the purpose of which is to normalize by rows and columns the obtained inverse (6). In summary, the three pitfalls to be avoided in second-order factoring are: taking primary correlations from raw scores instead of the simple structure plot; using reference vector correlations instead of factor correlations; and, as dealt

with below, leaving second-order factors orthogonal instead of seeking simple structure for them also.

The Questionnaire Foundation

The number of primary personality factors in questionnaire responses is still not fully determined, but the recent studies of Guilford (24) and of the present writer (10, 11, 12) converge in indicating that (among adults) about fifteen will account for most of the variance in highly varied collections of items. In the case of the author's studies and those of his research associates, the close coordination maintained between questionnaire and observer behavior-rating factors brings additional confirmation, in that some fifteen factors have also been found in the latter (3, 12).

The fifteen primary personality factors on which the present study is based have since 1949 been embodied in the Sixteen Personality Factor Questionnaire (15), into which, for general utility, general intelligence has been built as the sixteenth factor. The factor structure of the 432 items (466 less 34 intelligence items) comprising the A, B, and C forms of this test has been verified in three studies, each experimentally, statistically, and rotationally independent (5, 10, 11). A wide breadth of behavior reference in the 16 P.F. items has been guaranteed by the historical survey (3) and by the concept of sampling the personality sphere made in the first analysis (5), as well as by examination of 2,002 further items in the two succeeding studies (10, 11).

Hypotheses as to what the fifteen factors may be which so persistently appear have been put forward (3, 15, 22) and debated. However, it is only recently, with the return of external validation results from the 16 P.F. Test (9, 14, 15, 34), in clinical, educational, and industrial use, that these concepts have been evaluated and sharpened. For example, the examination of the degree of inheritance of these factors made on their equivalent forms in the Junior Personality Quiz (13) with eleven-year-old twins, siblings reared apart, etc. (14), has thrown much light on their basic natures. From these studies, it appears that 2 of the 15 factors (*A* and *H*) are hereditary components of the cyclothyme-schizothyme temperament continuum; that others, e.g., surgency-desurgency (*F*), and sensitivity (*I*), are almost wholly environmentally determined; that the factors *C* and *G* correspond to the clinical concepts of ego strength and superego strength; that *E*, *L*, and Q_1 correspond to the concepts of dominance, paranoid trend, and radicalism, and that *O* and Q_4 characterize two relatively distinct components in anxiety. More precise discussion can be given to these meanings of primaries where it becomes relevant to the interpretation of the second-order factors among them found here.

Correlations Among Primaries

Determination of the cosines (correlations) among the fifteen primary factors was based on the third (10) of the above studies on the 16 P.F., which had benefited from the preceding studies in acquiring more highly loaded items, and, therefore, a clearer factor structure. The correlation matrix was of order 75×75 and consisted of 2,775 product-moment correlations on an N of 408 subjects (181 undergraduate men and women students and 227 Air Force young men in training). Although the 16 P.F. test was represented by only 75 variables it will be seen by reference to the studies indicated (10, 11) that these truly represent the factor structure of the whole test. The special procedure of *parcelled factor analysis*, by which the 163 highest loaded items in the test are parcelled into 75 short scales of two and three items each (because of the impossibility of handling a 163×163 matrix) has been described elsewhere (11). It gives the test as a whole the structure of the 75 parcels, and evaluates each item in the total 16 P.F. against the factors established in this framework.

For purposes of determining the angles among oblique factors with great exactitude, in order to enter a second-order factor analysis, this design of having each factor loaded on only 5 out of 75 variables is particularly valuable. For the even distribution (each factor represented by the same number of variables) and the great preponderance of variables in every hyperplane (70 out of 75 is possible, and was closely approached in our

Table 1

Cosines Among Reference Vectors
Primary Personality Vectors

	(A)	(C)	(E)	(F)	(G)	(H)	(I)	(L)	(M)	(N)	(O)	(Q_1)	(Q_2)	(Q_3)	(Q_4)
(A)	1.01														
(C)	.13	1.00													
(E)	.07	.06	1.00												
(F)	−.01	.05	.00	.99											
(G)	.00	.04	−.02	.05	1.00										
(H)	−.36	−.12	−.18	−.26	.00	1.00									
(I)	−.29	−.04	−.06	.01	.00	.14	1.01								
(L)	−.01	−.01	−.06	−.11	−.04	.26	−.04	1.00							
(M)	.03	.14	.06	.26	−.04	.05	−.14	−.07	.99						
(N)	.08	.01	−.11	−.12	.13	−.05	.26	−.05	−.02	1.01					
(O)	.02	.19	.00	−.04	.02	.13	−.11	−.08	−.06	.04	1.00				
(Q_1)	−.15	.05	−.31	−.02	.11	−.09	.02	.01	−.19	.00	.01	1.01			
(Q_2)	.07	−.01	.18	.22	.02	.03	−.11	−.01	−.11	.03	.08	−.14	1.01		
(Q_3)	.30	−.19	.12	.22	−.19	−.29	−.02	.11	−.04	−.22	.20	.00	.12	1.00	
(Q_4)	.04	.09	.02	.05	−.05	.01	.15	−.15	−.01	−.07	−.17	.03	.01	.23	.99

study) permits a very sure perception of hyperplanes and their angles. Actually the mean percentage of variables obtained in the average hyperplane (between ± .10 loadings), for all factors, came to 91 per cent, which is larger than in any other published or unpublished study known to the writer. Secondly, the determination of each hyperplane by as many as fourteen factors, i.e., in fourteen-space, precludes any serious distortion of angles from the composition of the particular cluster of variables which might happen in this study to define, i.e., to be highly loaded in, any given factor.

The simple structure was actually approximated with relatively little labor by applying Wrigley's quartimax method (28). This gives the nearest approach by an analytical method, and within a Procrustean orthogonal structure, to the eventual oblique structure. However, since this analytic method stops short of the best simple structure, by tending to leave too much variance in the centroid and through failing to follow the natural obliquities, rotation was finished "by hand." Graphical plots of each factor with every other (105 plots per rotation) were made. After four rotations of all factors from the quartimax po-

Table 2

Correlations Among Factors
Primary Personality Factors

	(A)	(C)	(E)	(F)	(G)	(H)	(I)	(L)	(M)	(N)	(O)	(Q_1)	(Q_2)	(Q_3)	(Q_4)
(A)															
(C)	−.19														
(E)	−.02	−.04													
(F)	.16	−.04	.16												
(G)	−.04	.04	−.05	−.02											
(H)	.15	.32	.21	.28	.01										
(I)	.30	−.15	−.06	−.12	.02	−.24									
(L)	.04	−.26	.00	.03	−.02	−.42	.15								
(M)	−.06	−.20	−.08	−.36	.05	−.28	.23	.15							
(N)	−.22	.17	.16	.15	−.08	.26	−.40	−.11	−.13						
(O)	.13	−.44	−.02	.03	−.07	−.40	.23	.35	.17	−.22					
(Q_1)	−.16	−.03	.31	−.03	−.11	.08	−.05	−.02	.18	.08	−.03				
(Q_2)	−.06	−.06	−.20	−.33	−.01	−.24	.19	.08	.28	−.17	.04	.11			
(Q_3)	−.32	.47	−.03	−.14	.18	.42	−.24	−.39	−.08	.33	−.52	.04	−.11		
(Q_4)	.03	−.31	−.01	.02	−.02	−.28	−.03	.33	.09	−.05	.38	−.04	.03	−.42	

sition, a very clear structure was obtained. A fifth "polishing" rotation was carried out to give the best possible determination of angles. The mean hyperplane content indicated above, 91 per cent, was thus obtained, and the significances of the simple structure determinations, tested by Bargman's formula (1), was found to be beyond $p = .001$ for each of the fifteen factors.

The cosine matrix, $C = \lambda'\lambda$, among the reference vectors at this point is given in Table 1 and its inversion and normalization to the correlations among the factors, $K = DC^{-1}D$, in Table 2. It will be observed that as usual the intercorrelations among the factors are higher than among the vectors.

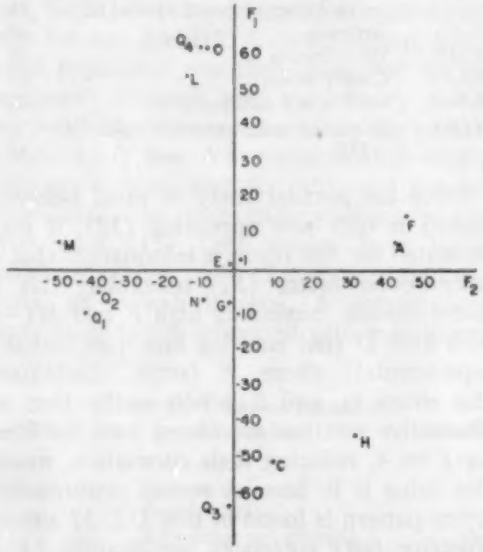

Fig. 1. Simple structure in the second-order realm: F_1 plotted against F_2.

The Second Order Factors

On the factor extraction from Table 2, carried out by the centroid method, Tucker's criterion (6) indicated four or five factors. Five were extracted, but subsequent rotation clearly collapsed one of them. One of the sources of confusion in second-order factor solutions, mentioned above, is that experimenters who rotate primaries for *oblique* simple structure nevertheless sometimes adopt *orthogonal*, unrotated, centroid interpretations for second-order factors. The well-known reasons for oblique simple structure in primaries (6, 30) actually apply just as forcibly to second-order factors. Simple structure may here be harder to determine without ambiguity, because of the fewer variables, but enough second-order studies now exist (12, 25, 30, 33) to show that simple structure commonly resides in the data and gives psychological meaning to factors thus uniquely determined. Accordingly, the second-order factors were rotated, and, after seven rotations, the structure shown in Table 3 was found. From the typical graph in Fig. 1 it will be evident that this is a very clear structure and practically unimprovable. (Possibly a 5° shift could still be made on each factor.) The number of variables in the hyperplane—

Table 3

Rotated Second-Order Factor Matrix

Primary factors	F_1	F_2	F_3	F_4
A	+.07	+.45	+.56	−.10
C	−.53	+.10	−.25	−.10
E	+.02	−.01	+.04	+.50
F	+.17	+.48	+.04	+.21
G	−.09	−.01	−.08	−.23
H	−.46	+.33	−.03	+.23
I	+.03	+.01	+.53	−.20
L	+.54	−.13	−.05	−.09
M	+.07	−.51	+.13	+.04
N	−.08	−.09	−.48	+.34
O	+.62	−.07	+.16	−.05
Q_1	−.12	−.42	+.05	+.55
Q_2	−.06	−.38	+.17	−.08
Q_3	−.66	−.01	−.29	−.02
Q_4	+.62	−.09	−.09	+.00

Note.—The unrotated matrix (the quartimaxed matrix, V_{oa}, from which the transformation matrix operates) and the transformation matrix by which this table was obtained have been deposited with the American Documentation Institute. Order Document No. 4969 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Table 4

Correlations of Second-Order Factors
(Simple structure after seven rotations)

	C^{-1}			
Factor	1	2	3	4
1	1.02			
2	−.02	1.21		
3	.15	−.29	1.09	
4	.02	.43	−.13	1.14

slightly more than half—is about that found in most satisfactory first-order structures.

The intercorrelations of second-order factors ($D(\lambda'\lambda)^{-1}D$), derived from the transformation (Lambda) matrix are shown in Table 4.

The chief indicators of the nature of Factors 1 through 4 in Table 4 will now be set out verbally, a contingent title being given to each factor.

Symbol	*Factor 1. Anxiety-vs.-Dynamic Integration*	Loading
$Q_3(-)$	Lack of will control[1]	.66
$O(+)$	Insecurity or free-floating anxiety	.62
$Q_4(+)$	Nervous tensions (somatic anxiety)	.62
$L(+)$	Paranoid trend	.54
$C(-)$	Lack of ego strength	.53

This factor has clearly brought together the "pathological" primaries. Before turning to interpretation one may point out that the evidence on second-order factors among behavior-rated personality primaries, which is published synchronously (12) with this article, shows a factor similar in all respects except that Factor I, called "Premsia" (12), is included there. Three theories compete for acceptance here: first, that it is a factor of general anxiety; second, a factor of general neuroticism; and third, that it is the different form of personality disorganization associated with psychoticism (20): see also Factor U.I. 25 in (7). The last has too many speculative elements and the first is somewhat weakened by the absence of the third main form of anxiety—desurgency ($F-$) or anxiety depression. On the other hand, it is striking that this set of five primary factors is identical with that found to correlate substantially with (and *only* with) the objective test factor (7) of Anxiety (indexed U.I. 24). (Both patterns also show some low H factor.) Nevertheless, considering the immediate nature of the above primary factors alone, one might incline to an interpretation of the second-order factor as Dynamic Integration-vs.-Neuroticism. There stands in the way of this interpretation Eysenck's interpretation of U.I. 23 as *the* neuroticism factor (20), an obstruction qualified somewhat by our reinterpretation of U.I. 23 as reduced neural reserves, rather than neurotic conflict per se. On the other hand, U.I. 24, with which the present second-order factor apparently identifies itself, is better described as anxiety than as conflict. Accordingly, since reliable description is more important than interpretation at this stage, it would seem best to compromise by naming the present second-order questionnaire factor *Dynamic Integration-vs.-Anxiety*.

[1] The titles used are those of the 16 P.F. Handbook (15). To make easy, consistent reading, all loadings have been made positive, by reflection, in these factor lists. The reflection is indicated by the sign on the left and the change of title to the opposite pole, e.g., "Lack of. . . ."

Symbol	*Factor 2. Extraversion-vs.-Introversion*	Loading
$M(-)$	Practical concern or praxernia (12) (vs. Bohemianism or alaxia [12])	.51
$F(+)$	Surgency	.48
$A(+)$	Cyclothymia	.45
$Q_1(-)$	Conservatism	.42
$Q_2(-)$	Lack of self-sufficiency	.38
$H(+)$	Adventurous cyclothymia or parmia (12)	.33

Since the parallel study in rated behavior factors is only now appearing (12), it may be stated for the reader's information that a second-order factor (F_4) resembling this is found therein, combining high F and $M(-)$ with high D (not entering into questionnaire experiments). There, E factor, dominance, also enters in, and it is noteworthy that an alternative rotation considered here for Factor 2 on 4, reducing their correlation, would also bring it in here. A second confirmation of the pattern is found in that U.I. 32 among objective tests correlates significantly ($p = .01$) with Factors F, A, Q_2, H and E; though M and Q_1 are not as significant. These are thus, once more, quite independent indications of the same pattern.

This pattern has pulled together two primary factors which have been long recognized as cyclothyme-schizothyme in character, but it emphatically adds others which might go under the broader concept of extraversion—such as surgency, dominance, practical concern and lack of self-sufficiency. Along with conservatism these point to a person closely identified with the group life—"a good fellow"—as opposed to the independent recluse. The relatively low loadings of the more hereditarily determined H and A fac-

tors (14) suggest that Factor 2 is an environmentally determined pattern. There now seems ground for recognizing extraversion-introversion as something more than a correlation cluster or surface trait, and perhaps for specializing the term for this measurable second-order factor.

Symbol	*Factor 3. Cyclothyme-vs.-Schizothyme Constitution*	*Loading*
$A(+)$	Cyclothymia	.56
$I(+)$	Emotional sensitivity or premsia (12)	.53
$N(-)$	Lack of sophistication	.48
$Q_3(-)$	Lack of will control	.29
$C(-)$	Lack of ego strength	.25

Again, a somewhat similar factor is found in the second-order among behavior-rated factors, in No. 5 (12), which links cyclothymia with low ego strength. A constitutional, body-build hypothesis was there suggested for experiment. Although only preliminary results are available it seems also to be the pattern (high *A, I,* low *N*) distinguishing manic-depressives from normals (15), and has fair resemblance to that distinguishing women from men (15) (but women have higher manic-depression and lower schizophrenia rates, as is well known). A constitutional body-build factor connected with cyclothymia is hypothesized.

Symbol	*Factor 4. Unbroken Success-vs.-Frustration*	*Loading*
$Q_1(+)$	Radicalism	.55
$E(+)$	Dominance	.50
$N(+)$	Sophistication	.34
$G(-)$	Low super ego strength	.23

Again it is noteworthy that a second-order factor (Factor 2) among observer-rated behavior factors (12) presents a pattern agreeing excellently with this in the variables common to the two studies (*E, N,* and *G*). The present study adds Q_1 and the objective behavior study adds high intelligence. Possible hypotheses are "easy personal success" occasioned by good intelligence or good educational opportunity, or "social status," which is known to be correlated with intelligence and dominance (2), and could well be correlated with sophistication and radicalism, in young men. Incidentally, the reader is reminded of the evidence that "radicalism," as a factor, is more than an accident of political belief, and constitutes a general temperamental source trait (19, 29).

Doubtless a generous native endowment in abilities plays a large causative role in this syndrome. One is reminded of a famous ambassador's comment on a British Secretary of State, known for his radical innovations, dominance of the world scene, sophistication, and debauchery (low *G!*): "his capacity is on so high a level that he has no need to give up his pleasures in the slightest degree for any cause" (17, p. 478). But in the life-record factor (12) the intelligence loading is scarcely high enough to make it the essential root, and the loadings are similar to those for social status. Consequently it is safer to consider this as the organization within personality brought about by unbroken success, from whatever cause, as opposed to frustration and intellectual bafflement. Among students it should correlate negatively with age.

Summary

1. A questionnaire—the 16 P.F.—broadly based on the personality sphere and over 2,000 items, had its factor structure redetermined on a sample of 408 young men and women, with particular attention to exact rotation to the best simple structure.

2. The correlations thus determined, i.e., by inversion of the reference vector matrix, among the 15 confirmed primaries (intelligence was omitted) factored to four second-order, simple-structure factors.

3. These four factors have resemblances, ranging from fair to excellent, to four second-order factors found from life-record (behavior rating) primaries, and two of them have closely similar profiles to the correlations found between questionnaire primaries and two objective personality test factors. The latter seems to confirm the surmise, made on general principles, that objective test factors would tend to be on a different level from rating and questionnaire factors, and, generally, to be second-order to them.

4. The factors have been contingently named *Dynamic Integration-vs.-Anxiety* (Objective test factor: U.I. 24); *Extraversion-vs.-Introversion,* (Objective test factor: U.I. 32); *Cyclothyme-vs.-Schizothyme Constitution;* and *Unbroken Success-vs.-Frustration.*

There is a significant positive correlation here of the second and fourth, as also was found in behavior rating second-order factors (12). Definite hypotheses are proposed for their further experimental investigation. As a part of such investigation it is desirable to score individuals on the 16 P.F. not only for primaries but also for these four second-order factors (from the above loadings) to see if higher correlations with a specific criterion are associated with second-order than with first-order factors. *Dynamic Integration* scores should be of particular interest to the clinician.

Received January 24, 1956.

References

1. Bargman, R. Significanzuntersuchungen der Einfachen Struktur in der Faktoren-analyse. *Mitteil. Math. Statist.*, Sonderdruk. Physica-Verlag. Wurzburg, 1954.
2. Cattell, R. B. The cultural functions of social stratification: 1. Regarding the genetic bases of society. *J. soc. Psychol.*, 1945, **21**, 3–23.
3. Cattell, R. B. *The description and measurement of personality*. Yonkers, N. Y.: World Book Co., 1946.
4. Cattell, R. B. Oblique, second-order and co-operative factors in personality analysis. *J. gen. Psychol.*, 1947, **36**, 3–22.
5. Cattell, R. B. The main personality factors in questionnaire, self estimate material. *J. soc. Psychol.*, 1950, **31**, 3–38.
6. Cattell, R. B. *Factor analysis*. New York: Harper, 1952.
7. Cattell, R. B. The principal replicated factors discovered in objective personality test. *J. abnorm. soc. Psychol.*, 1955, **50**, 291–314.
8. Cattell, R. B. Psychiatric screening of flying personnel. Personality structure in objective tests —A study of 1000 Air Force students in basic pilot training. *USAF Sch. Aviat. Med. Proj. Rep.*, 1955, Proj. No. 21-0202-007 (Rep. No. 9).
9. Cattell, R. B. Occupational profiles on the I.P.A.T. 16 Personality Factor Questionnaire. *Occup. Psychol.*, 1956, **30**, 1–10.
10. Cattell, R. B. Validation and intensification of the sixteen personality factor questionnaire. *J. clin. Psychol.*, in press.
11. Cattell, R. B. A shortened "Basic English" version (Form C) of the 16 P.F. Questionnaire. *J. soc. Psychol.*, in press.
12. Cattell, R. B. *Personality and motivation structure and measurement*. Yonkers, N. Y.: World Book Co., in press.
13. Cattell, R. B., Beloff, J., Flint, D., & Gruen, W. The Junior Personality Quiz. Champaign, Ill.: Institute for Personality and Ability Testing, 1602 Coronado Dr., 1954.
14. Cattell, R. B., Blewett, D. B., & Beloff, J. R. The inheritance of personality. A multiple variance analysis determination of approximate nature-nurture ratios for primary personality factors in Q-data. *Amer. J. hum. Genet.*, 1955, **7**, 122–146.
15. Cattell, R. B., Saunders, D. R., & Stice, G. F. The Sixteen Personality Factor Questionnaire. Champaign, Ill.: I.P.A.T., 1602 Coronado Dr., 1949.
16. Cattell, R. B., & Stice, G. F. Four formulae for selecting leaders on the basis of personality. *Hum. Relat.*, 1954, **7**, 493–507.
17. Churchill, W. S. *Marlborough, his life and times*. Vol. IV. London: Harrap, 1938.
18. Dexter, E. S. Personality traits related to conservatism and radicalism. *Charact. & Pers.*, 1939, **7**, 230–237.
19. Drevdahl, J. F. Factors of importance for creativity. *J. clin. Psychol.*, 1956, **12**, 21–26.
20. Eysenck, H. J. *The structure of human personality*. London: Methuen, 1953.
21. Fisher, R. A. *Social selection of human fertility*. New York: Oxford Univer. Press, 1932.
22. French, J. The description of personality measurements in terms of rotated factors. Princeton, N. J.: Educational Testing Service, 1953.
23. Guilford, J. P., & Guilford, Ruth B. Personality factors D, R, T and A. *J. abnorm. soc. Psychol.*, 1939, **34**, 1–12.
24. Guilford, J. P., & Zimmerman, W. S. The Guilford-Zimmerman Temperament Survey. Beverly Hills, Calif.: Sheridan Supply Co., 1949.
25. Hsü, E. H. On the correlation between a variable and its super-factor. *J. Psychol.*, 1946, **22**, 89–92.
26. Lentz, F. F. Generality and specificity of conservatism-radicalism. *J. educ. Psychol.*, 1938, **29**, 540–546.
27. Lovell, C. A study of the factor structure of thirteen personality variables. *Educ. psychol. Measmt*, 1945, **5**, 335–350.
28. Neuhaus, G. O., & Wrigley, C. The quartimax method: an analytical approach to orthogonal simple structure. *Brit. J. statist. Psychol.*, 1954, **7**, 81–91.
29. Suhr. V. W. The Cattell 16 P.F. as a prognosticator of accident susceptibility. *Proc. Iowa Acad. Sci.*, 1953, **60**, 553–561.
30. Thurstone, L. L. *Multiple-factor analysis*. Chicago: Univer. of Chicago Press, 1947.
31. Thurstone, L. L. Thurstone Temperament Schedule. Chicago: Science Research Associates, 1949, 1950.
32. Thurstone, L. L., & Thurstone, Thelma G. Factorial studies of intelligence. *Psychometr. Monogr.*, 1941, No. 2.
33. Vernon, P. E. *The structure of human abilities*. London: Methuen, 1950.
34. Wright, S. Some psychological and physiological correlates of certain academic underachievers. Unpublished doctor's dissertation, Univer. of Chicago Library, 1955.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Factors, Response Bias, and the MMPI

Eugene H. Barnes
VA Hospital, Downey, Illinois

Certain deductions concerning the influence of response set on the MMPI were tested in a recent article by Barnes (1). These deductions were derived from Berg's deviation hypothesis (2) and other pertinent research. This hypothesis asserts: "Deviant response patterns tend to be general; hence those deviant behavior patterns which are significant for abnormality and thus regarded as symptoms, are associated with other deviant response patterns which are in noncritical areas of behavior and which are not regarded as symptoms of personality aberration" (2, p. 62). Since the scoring system of the individual form of the MMPI provides that only those items answered in the infrequent direction are recorded, the MMPI provided a convenient source of deviant responses as defined by Berg. In his article Barnes (1) presented correlations of the MMPI scales with the number of items answered atypically true and with the number answered atypically false. He found the tendency to deviate by answering false to be more highly correlated with the scales of the neurotic triad (*Hs, D, Hy*) and the tendency to deviate by answering true to be more highly correlated with the scales of the psychotic triad (*Pa, Pt, Sc*). Further, the two tendencies were not significantly correlated with one another.

A factor analysis of the MMPI scales based on the scores of 110 neuropsychiatric patients was reported by Wheeler, Little, and Lehner (6) in which two major factors and one minor factor were identified. Factor I, the "psychotic" factor, had its maximal loading on *Pt* and *Sc*. This factor also had a substantial negative loading on *K*. Factor II, the "neurotic" factor, had its maximal loadings on the scales of the neurotic triad.

The correlations reported by Barnes (1) are sufficiently similar to the factor loadings on the two major factors reported by Wheeler, Little, and Lehner (6) to suggest that the sets may be empirical correlates of the rationally derived factors. Specifically, it is hypothesized that, if included in the correlation matrix, the vector representing number of true answers would be approximately collinear with Factor I, while the number of atypical false answers would be approximately collinear with Factor II.

Procedure

A direct test of this hypothesis does not appear possible since different samples are involved. An indirect method is to assume the existence of a test whose vector is collinear with the reference vector, calculate the correlations of each test in the matrix with the hypothetical test, and compare these calculated estimates with the obtained correlations of the test under investigation. In order to make these calculations it is necessary only to know the communalities of the tests in the original matrix, the cosine of the angle each makes with the reference vector, and the communality of the hypothetical test (5). Refined estimates of the communalities of the tests in the matrix are provided by the factor analysis. The cosine of the angle each test makes with the reference vector is simply its factor loading divided by the square root of the communality. The communality of the hypothetical test must be estimated. In this predicament it seems logical to use the same approximation recommended for such estimates in the original factor problem, i.e., the highest correlation obtained between the suspected collinear test and any of the other

Table 1

Correlations of Atypical Responses and Hypothetical Tests with Scales of MMPI

Scale	Atypically false	Hypothetical Test II	Atypically true	Hypothetical Test I
L	.37	—	−.24	—
F	.43	—	.66	.64
K	.06	—	−.69	−.67
Hs	.56	.63	.52	.50
D	.73	.55	.34	.46
Hy	.59	.72	−.09	—
Pd	.28	—	.68	.67
Mf	.21	—	.28	.36
Pa	.37	—	.52	.60
Pt	.31	—	.84	.89
Sc	.36	—	.90	.88
Ma	.21	—	.50	.59

Table 2

Communalities and Reliabilities of MMPI Scales

	Communalities	Reliabilities	
Scale	Wheeler, Little, & Lehner (6)	Rosen (4)	Holzberg & Alessi (3)
L	.234	.62	.85
F	.700	.81	.93
K	.643	.65	—
Hs	.873	.85	.67
D	.741	.80	.80
Hy	.771	.88	.87
Pd	.551	.88	.52
Mf	.338	.64	.76
Pa	.780	.75	.78
Pt	.944	.80	.72
Sc	.878	.83	.89
Ma	.494	.56	.58

tests in the matrix. In order to obtain estimated correlations using this approximation it is merely necessary to compute the scalar product of the two vectors, which turns out algebraically to be the square root of the estimated communality multiplied by the factor loading of the test with which its correlation is to be inferred. The communality of the test suspected to be collinear with Factor I (atypical true answers) is estimated at .90, its correlation with *Sc*. The communality of atypical false answers is estimated at .73, its correlation with *D*.

Results

The correlations expected for the hypothetical collinear tests as compared with the correlations obtained by Barnes (1) are presented in Table 1. It will be noted that estimates are not made for all tests on both factors. This is because factor loadings were not reported on these tests by Wheeler, Little, and Lehner (6). It is presumed that these omitted factor loadings were not considered significant. This would be consistent with the low predicted correlations.

It is obvious that the predicted correlations for a hypothetical test collinear with Factor I are in very close agreement with the experimental results of correlating the number of atypically true answers with the scales of the MMPI. The agreement of the predicted and obtained correlations for the number of atypically false answers is not so close, nor yet does it seem so divergent as to refute the hypothesis. Unfortunately there does not seem to be any significance test which can be applied to the kind of data presented. The meaning of the results cannot be evaluated, therefore, with the precision which is desirable.

Discussion

The correlations between the number of items atypically answered true and the scales of the MMPI are cogently similar to the correlations predicted for the hypothetical test, despite the lack of the significance test. There would appear to be little risk involved in accepting this measure as a "pure factor test" of the "psychotic factor." Although the predicted correlations for the number of atypical false answers do not have such a striking similarity, it is clear that, if included in the test matrix, this measure would have a heavy loading on the "neurotic factor." For this reason scores based on these response sets should, in themselves, be of considerable interest. It should be remarked, however, that a measure of the "psychotic factor" cannot be accepted as a measure of psychotic tendencies without further experimental evidence. A similar and more emphatic caution applies to the "neurotic factor," and the tendency atypically to answer false.

In addition to their interest as measures, it

would seem that they might have another value. A comparison of the communalities reported by Wheeler, Little, and Lehner (6) with reliability coefficients reported by Rosen (4) and by Holzberg and Alessi (3) suggests that a degree of specificity remains in some of the scales. This comparison appears in Table 2. Since these two factors accounted for most of the communality of the scales of the MMPI and, since the two response sets appear to be relatively pure measures of the factors, it would appear that some adjustment of the scoring should be possible to reveal the remaining specificity. Such an adjustment should considerably increase the diagnostic precision of the MMPI. It is not clear to the writer exactly how this adjustment could be accomplished, but it seems inescapable that some allowance for response set will be necessary in the interpretation of the MMPI.

It is possible that the method of analysis used in this study may recommend itself to other investigators. The technique may be of value to the test constructor seeking to develop "factor pure" tests as a check of sorts on his success. Further, it might be used as a method of selecting measures to be included in a matrix for factor analysis. It would seem particularly appropriate if a previous analysis has given suggestive but somewhat equivocal results.

Summary

Two major factors, a "neurotic" factor and a "psychotic" factor, were identified by Wheeler, Little, and Lehner (6) in a factor analysis of the scales of the MMPI. Hypothetical tests collinear with these factors were postulated and the correlations of these hypothetical tests were computed for each of the scales in the correlation matrix for which factor loadings were reported. These correlations were compared with those obtained by correlating the number of atypical true answers to the MMPI with the scales of the MMPI and those obtained by correlating the number of atypical false answers with the scales of the MMPI. It was concluded that the number of atypical true answers is a "pure factor test" of the "psychotic factor" and that the number of atypical false answers has a heavy loading on the "neurotic factor." It was also concluded that some allowance should be made in the interpretation of the MMPI to allow for the influence of response set.

Received February 6, 1956.

References

1. Barnes, E. H. Response bias and the MMPI. *J. consult. Psychol.*, 1956, **20**, 371–374.
2. Berg, I. A. Response bias and personality: The deviation hypothesis. *J. Psychol.*, 1955, **40**, 61–72.
3. Holzberg, J. D., & Alessi, S. Reliability of the shortened MMPI. *J consult. Psychol.*, 1949, **13**, 288–292.
4. Rosen, A. Test-retest stability of MMPI scales for a psychiatric population. *J. consult. Psychol.*, 1953, **17**, 217–221.
5. Thurstone, L. L. *Multiple-factor analysis*. Chicago: Univer. of Chicago Press, 1947.
6. Wheeler, W. M., Little, K. B., & Lehner, G. F. J. The internal structure of the MMPI. *J. consult. Psychol.*, 1951, **15**, 134–141.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Personality of Judges and Their Agreement with Experts in Judging Clinical Case Histories

A. W. Bendig[1]
University of Pittsburgh

Large individual differences among judges in assessing clinical case histories suggest that variables such as intellectual ability, educational level, interests, and personality traits may influence their judgments. Such factors may be more important than the formal characteristics of the judgmental situation: the type of clinical material judged or the method of recording the judgments. The present study was designed as a preliminary investigation of the possible relation between certain objective personality traits and judgmental ability.

Undergraduate college *Ss* ($N = 96$) ranked for global adjustment level 10 short clinical case histories with approximately half of the *Ss* ranking a set of 10 cases of adult clinic patients and the other half ranking 10 case histories of disturbed children. The *Ss* also completed the 15-scale Edwards Personal Preference Schedule (1). The rankings of each *S* were correlated with similar rankings of the same cases by 6 psychiatrists (adult cases) or 7 clinical child psychologists (child cases) and the *Ss* dichotomized on the basis of the magnitude of their agreement with these experts. The *Ss* were further divided into two replicated groups and the high and low agreement *Ss* of Group I ($N = 56$) compared on each of the 15 scales of the PPS by two criterion (case type and agreement group) analysis-of-variance procedures. Two of the PPS scales, Achievement and Exhibition, significantly (.01 level or better) discriminated between the high and low agreement groups, while the interaction of case type and agreement group was significant (.05 level) with the Order scale. The high agreement group was higher on both the Achievement and Exhibition scales, while the high agreement *Ss* ranking the adult cases and the low agreement *Ss* ranking the child cases scored higher on the Order scale. Similar analyses of variance were computed on these three scales for the replication Group II ($N = 40$) and none of the resulting F ratios approached significance. Again the high agreement *Ss* showed high mean scores on the Achievement and Exhibition scales, but the differences were not significant.

The obvious conclusion of this study is that there are no consistent relationships between the scores of naive *Ss* on the scales of Edwards' PPS and their agreement with experts in judging clinical case histories. However, this does not exclude the possibility that other personality measures might show some correlation with judgmental ability.

Brief Report.
Received July 6, 1956.

[1] An extended report of this study may be obtained without charge from A. W. Bendig, Dept. of Psychology, University of Pittsburgh, Pittsburgh 13, Pa., or for a fee from the American Documentation Institute. Order Document No. 5039 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Reference

1. Edwards, A. L. *Manual for the Edwards Personal Preference Schedule* New York: Psychological Corp., 1954.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Comparison of Two Methods of Electroshock in Their Effect on Cognitive Functions[1]

Barbara S. Kendall, Warren B. Mills,
Washington University School of Medicine
and Thomas Thale
St. Louis University School of Medicine

Electroconvulsive therapy (ECT) commonly produces temporary confusion and disturbance of attention and of memory in the treated patient (6), or, as Stone (10) points out, a general impairment of cognitive functions. These unfortunate effects may accumulate during the treatment period, but tend to disappear within two or three weeks after the final shock. It has usually been assumed that ECT by the Reiter apparatus, because a lesser amount of current is necessary to produce a seizure, results in less disturbance to cognitive functions than does the conventional method (1, 2, 12, 13). Yet there have been no well-controlled studies dealing with measured differences in the effect on "test ability" of these two methods.

The aim of the present study is to test the following hypotheses:

1. For the same therapeutic result ECT by the Reiter method causes a lesser amount of impairment of cognitive functions, demonstrable at the end of the shock series, than does ECT by the conventional method.
2. Patients shocked by the Reiter method show less residual cognitive disturbance two weeks after the final convulsion than do patients treated by the conventional method.
3. Two weeks after the last convulsion, score gains in cognitive tasks over preshock performance are greater and losses are less for patients treated by the Reiter method than for those treated by the conventional method.

[1] The authors acknowledge their indebtedness to Dr. Goldine C. Gleser for help in planning the statistical treatment of the data and to Dr. Laverne C. Johnson for reading the manuscript and for making valuable suggestions. This research was done at the Malcolm A. Bliss Psychopathic Hospital.

Method

Amount and type of treatment. Two groups of patients treated at the Malcolm A. Bliss Hospital are compared with reference to cognitive changes induced by ECT; one group, in which all were treated with the Reiter type 46 E machine (which delivers pulsating "unidirectional" current), will be referred to as the "Reiter group"; the "Conventional group" consists of those receiving treatment with the Lektra machine model 302 (which puts out a single burst of unmodulated 60-cycle sine wave alternating current). In order to insure uniformity, all patients were treated with a bitemporal placement of the electrodes, and no patient was presedated. This procedure was the usual one followed by this hospital in administering ECT. Treatments were scheduled three times a week until a full course had been run. Since it had been the experience in this institution that 20 convulsive reactions on the Reiter machine were usually required to induce approximately the same therapeutic result as that induced by 12 convulsions on the conventional type machine, each patient in the Reiter group received 20 and each one in the Conventional group 12 treatments. In a few instances, because of mechanical and other difficulties, single interruptions of the treatment rhythm resulting in intervals of four to seven days (instead of the expected

two to three days) were unavoidable. These occurred only once for any one patient, involved four individuals in the Conventional and nine in the Reiter group, and tended to be longer in the latter group. Later it will be shown that these single interruptions did not alleviate the detrimental effect of ECT. One patient in the Reiter group received her sixteenth treatment on a Medcraft (conventional) machine.

Sampling. The two groups of 20 patients each were matched for age, education, and intelligence. The mean age of the Reiter group was 30.40 years with a range of 21 to 52 years and an *SD* of 9.15; the mean age of the Conventional group was 34.35 years (range 20 to 59 years and *SD* of 9.59). For the Reiter group the mean number of years of education was 9.70 with a range of 6.50 to 15.00 and an *SD* of 2.51; for the Conventional group, it was 10.08 with a range of 4.00 to 16.50 and an *SD* of 2.89. The Wechsler-Bellevue Vocabulary was used to obtain a stable measure which would approximate original intellectual endowment. Form I was given before treatment and Form II two weeks after the close of treatment; whichever rating was higher was used for this score. The mean weighted score of the Reiter group was 10.70 points with a range of 7 to 16 and an *SD* of 2.41; that for the Conventional group was 10.75 points, with a range of 4 to 16 and an *SD* of 2.79.

There were 9 men and 11 women in the Reiter group and 8 men and 12 women in the Conventional group. Psychiatric diagnoses for the Reiter group were distributed as follows: schizophrenic reaction, 15 (of which 4 were of schizo-affective type); paranoid state, 2; involutional psychotic reaction, 2; and manic-depressive reaction, depressive type, 1. For the Conventional group they were: schizophrenic reaction, 16 (of which 2 were of schizo-affective type); paranoid state, 1; involutional psychotic reaction, 1; manic-depressive reaction, depressive type, 1; and depressive reaction, 1. None of the patients had received any type of convulsive treatment within the previous year; no case with known brain damage was included.

It was not possible to match the two groups for severity of illness, and selection from this point of view was necessarily at random. However, certain rough checks may be applied. Duration of illness is not significantly different in the two groups: 9 patients in the Reiter and 8 in the Conventional group had been ill for one year or less; comparable figures when the criterion is two years are 12 and 10. An index of Wechsler-Bellevue scatter which Pascal and Zeaman (8) have shown to differentiate between degrees of severity of psychiatric illness yields means of 9.35 for the Reiter and 8.06 for the Conventional group; the difference is not significant (t equals 1.01).

Psychological tests. In order to measure the amount of postshock confusion two short test batteries were set up, to be administered on the day of the first treatment, on the day when half the course of electroshock was completed (in the case of the Reiter group on the tenth treatment day and for the Conventional group on the sixth day), and, finally, on the last day of treatment.

The first battery, consisting of short tests of Picture Absurdities (7), Substitution (7), Cancellation (Whipple), and Cube Imitation (Knox), was invariably given 30 minutes after the convulsion, and will be referred to as the Morning Battery. The second, used (with few exceptions) three to three and a half hours after the convulsion, will be called the Noon Battery; this included naming the months in reverse, subtracting serial 7's (9), and serial 3's, naming colors (9), duplicating from memory patterns of colored blocks (9), and identifying from a pack of 15 cards 5 which carried words just previously used as stimulus words in an opposites test (5). Two patients in the Conventional group, one an aggressive chronic paranoid and the other a catatonic in her first episode, refused to cooperate in the Noon Battery on the first shock day, but gave adequate cooperation three hours after the second treatment. Where possible, alternate forms of the tests of these batteries were arranged for use at the second and third sessions. No change seemed practicable in Months Reversed, Picture Absurdities, Cube Imitation, Color Naming, or the Ruesch Blocks; but slight changes were made by shifting the key in Substitution, crossing out a different letter in Cancellation, and beginning with a different number in Serial 7's and Serial 3's; for the Recognition test the subject was asked to find opposites for a new set of words. Thus, although transfer of training could not be eliminated, the influence of previous practice was minimized for half of the tasks without introducing, except in the case of Recognition, a real uncertainty about equivalence of difficulty.

In order to have a measure of the rate of recovery from the temporary intellectual inefficiency which is the immediate cumulative effect of ECT, the Wechsler Memory Scale Form I was administered to all patients immediately after the Noon Battery on the final shock day, and this was followed approxi-

mately two weeks later with Form II of the same scale. Longer-term losses and gains in intellectual efficiency which resulted from the treatment were evaluated by means of the Wechsler-Bellevue Scale. When possible, Form I was given before the first shock, and Form II 13 to 15 days after the last shock.

Results

Measures of Postshock Confusion

In the analysis of results *T* scores were computed, according to the method of McCall, for each subtest of the Morning and Noon Batteries, and these were based (with one exception) on all three performances of both groups in the same or equivalent tasks. The Recognition test was given after the first and final shocks only, and, consequently, 80 scores, instead of 120, were used in the calculations. It would not have been useful to score the subtests of the battery by already established norms (where these existed) because the scores earned after ECT tended to accumulate at the lower end of each distribution to an extent which precluded effective differentiation. It was also undesirable to normalize for each administration separately because the scores would have failed to show the extent of the changes which occurred in the course of treatment. The method chosen, i.e., the normalizing of scores of each subtest for all three administrations, does not represent an attempt to establish standards of performance, but is rather a means of providing a basis for comparisons between Reiter and Conventional groups and among the three administrations of the subtests.

In Table 1 mean scores for each group in each subtest administered on the first day of treatment are compared with mean scores for those given on the last day of treatment, and a cross comparison is made of the differences between gains and losses shown by both groups. Means were computed for midcourse measures, but, since they add little to the analysis, are not included in the table.

Changes induced by ECT. Test-retest changes are considered separately for each experimental group. Decrements significant at or beyond the 2% level of confidence occur for both groups in Substitution and Serial 7's, and at or beyond the 5% level in Cancellation. For one or the other group losses significant at or beyond the 1% level are found in Picture Absurdities, Serial 3's, and Recognition, and at the 5% level in Cube Imitation and Months Reversed. Thus 8 of the 10 subtests are shown to be more or less sensitive to the cognitive changes induced by ECT. In the case of Color Naming, however, the effect of previous practice appears to cancel out the damaging effect of ECT; and Block Design actually shows a gain, though for the Reiter only, which is significant at the 5% level.

Test of Hypothesis I. The difference in gain and loss from first to final shock is expressed as the average change for Reiter minus the average change for Conventional. A comparison of the two groups with regard to these score decrements or increments yields only one significant difference, that for Recognition. A possible reason why this test of memory and alertness made the differentiation where the other tests of postshock confusion failed to do so is that transfer of training is more successfully minimized for this task than for any of the others, partly because a new set of stimulus words is used in the retest, and partly because the midcourse retest is omitted. It may be that for the other subtests the effect of previous practice offset in some degree the effect of confusion, but in a test of new learning the Reiter group was better able to demonstrate superiority.

Memory Scale

Changes in first two weeks of recovery period. In contrast to the experience with tests used to measure changes taking place during the course of treatment, practice effects in the case of the Wechsler Memory Scale, which is used as a measure of recovery, tend to inflate the retest scores and make the differences artificially greater. In Table 2 the mean Memory Quotients on the last day of ECT are compared with results of a second administration two weeks later. For both groups the test-retest gains are significant beyond the 1% level of confidence.

Further test of Hypothesis I. While the majority of the tests of postshock confusion

Table 1

Measures of Postshock Confusion: Means, Sigmas, and Measures of Difference in Reiter (R) and Conventional (C) Groups (*T* Scores)

Test	First shock (1)		Final shock (3)		Diff.			
	Mean	Sigma	Mean	Sigma	1−3	*t*	$D_R - D_C$	*t*
Morning battery								
Picture Abs.								
R	52.60	6.04	46.20	9.35	6.40	2.86***	2.10	.63
C	52.10	7.30	47.80	9.28	4.30	1.76		
Substitution								
R	54.60	11.04	45.85	9.52	8.55	3.88***	1.25	.35
C	53.90	9.43	46.40	8.27	7.50	2.72**		
Cancellation								
R	52.30	9.23	47.25	9.71	5.05	2.12*	−.40	.13
C	53.05	9.34	47.60	9.91	5.45	2.75**		
Cube Imitation								
R	54.40	8.61	50.05	8.86	4.35	2.21*	.85	.23
C	49.40	10.45	45.90	9.33	3.50	1.13		
Noon battery								
Months Rev.								
R	53.75	10.79	48.15	11.50	5.60	2.32*	2.60	.92
C	50.95	9.28	47.95	8.64	3.00	2.02		
Serial 7's								
R	54.95	10.78	48.30	9.18	6.65	4.21***	1.65	.84
C	52.00	10.27	47.00	7.93	5.00	4.23***		
Serial 3's								
R	50.60	12.69	49.10	8.77	1.50	.59	−2.20	.79
C	51.35	9.63	47.65	8.56	3.70	3.29***		
Color Naming								
R	51.05	13.00	48.63	10.62	2.42	1.08	−.58	.17
C	51.15	10.25	48.15	8.26	3.00	1.18		
Blocks								
R	46.55	10.50	52.90	9.85	−6.35	2.42*	−3.75	1.15
C	47.55	9.43	50.15	3.44	−2.60	1.35		
Recognition								
R	51.00	6.98	53.60	9.26	−2.60	1.01	−9.80	2.72***
C	50.70	9.37	43.50	7.10	7.20	2.86***		

* Significant at the 5% level of confidence.
** Significant at the 2% level of confidence.
*** Significant at the 1% level of confidence.
Note.—Means are rounded. In all cases higher scores indicate more favorable performance.

fail to bring out clear-cut differences between the effects of the two methods of administering ECT, results for the Wechsler Scale on the final shock day give striking evidence that the Reiter group shows less cognitive disturbance. Significance beyond the 1% level of confidence is indicated.

A comparison of subtest scores reveals that the Reiter group surpasses the Conventional in several measures of new learning, i.e., in Logical Memory, Memory for Designs, and the "hard" items of Associate Learning; differences of 2.73, 1.90, and 1.05 are significant (in the first instance at the 1% level, and for the last two at the 5% level of confidence); *t* is, respectively, 2.90, 2.08, and 2.10. No other difference in subtest scores between the two groups (except in Orientation) approaches statistical significance.

Because the two groups are matched for intelligence, age, and education it is assumed that differences found at the first administration of the Memory Scale reflect differences in the amount of cognitive impairment induced by the ECT series. An attempt to match the groups for the memory factor by

Table 2

Measures of Recovery from Postshock Damage: Means and Sigmas in Reiter (R) and Conventional (C) Groups

Test	Pretreatment (I)		Final shock (I)		After 2 wks. (II)	
	Mean	Sigma	Mean	Sigma	Mean	Sigma
Wechsler MQ						
R			76.95	11.36	90.60	20.28
C			66.55	8.86	81.90	16.95
Wechsler-Bellevue IQ						
R(*N* 20)	98.85	12.68			95.05	13.26
C(*N* 16)	103.31	12.27			99.50	10.16
C(*N* 17)	103.47	11.92				
C(*N* 19)					96.42	13.30

	Differences and their significance							
	I−II	*t*	RI−CI	*t*	RII−CII	*t*	D_R-D_C	*t*
Wechsler MQ								
R	−13.65	3.47**	10.40	3.15**	8.70	1.44	−1.70	.32
C	−15.35	4.39**						
Wechsler-Bellevue IQ								
R	3.80	1.89	−4.62	1.11	−1.37	.31	−0.01	.01
C	3.81	2.19*						

* Significant at the 5% level of confidence.
** Significant at the 1% level of confidence.
Note.—The Wechsler-Bellevue data for the Conventional group are incomplete; one patient was mute and two had been given only four subtests prior to shock; a fourth refused the posttest Performance Scale. Differences are calculated for the varying numbers of patients who gave full records in the different situations.

using the scale as a pretest or as a posttest would have been illusory. The groups comprised a variety of psychiatric conditions, and the schizophrenic patients were in different stages of the illness from acute to chronic. Any pretest, other than vocabulary, would have fallen far short of representing inherent intellectual endowment. The remission of the psychosis brought about by ECT should be maximal immediately after the conclusion of the series, and its effect more nearly equal for all patients than at any other time, but this effect was essentially nullified for cognitive tests by the intellectual impairment which all the experimental patients suffered. Two weeks later, when the retest was made, some patients had begun to lose the benefits accruing from the treatment while others maintained their gains. In either situation the establishment of approximate equivalence for the two groups in a sensitive measure was not possible. Matching for the more stable attributes of vocabulary, age, and education, though not ideal, was the most that could be achieved.

Test of Hypothesis II. Two weeks after the last treatment the difference between the means of the two groups in the Memory Scale, though not greatly diminished, is no longer statistically significant, and the difference in the extent of test-retest changes, i.e., in mean gains or losses, is negligible. A wider spread of scores, shown in what is almost a doubling of the standard deviations, reflects a variability in immediate outcome. Stone (11) has found the Wechsler Memory Scale to be "remarkably sensitive" to the temporary intellectual impairment associated with ECT. It seems also to be sensitive to changes which, in those cases which derive no lasting benefit, are associated with a reactivation of the psychosis for which a temporary remission has been effected by ECT. Ten patients of the Reiter and 12 of the Conventional group had improved sufficiently to be slated for an early discharge. Among the remaining patients, three were signed out against advice, a few were discharged later, but others required long-term care. The mean gains for the "good" outcome group—20.00

—and for the "poor" outcome group—7.78 —differ by 12.22 points, and this difference is significant at the 2% level.

Effect of interrupted treatment. Since the amount of gain shown in the final administration of the Memory Scale might be thought of as a rough measure of the extent of impairment of brain function which resulted from the full course of ECT—impairment which was actually reversed in the two-week period following treatment—an attempt is made to use this measure to determine whether the patients whose courses of treatment were interrupted had suffered less temporary detrimental effect as a consequence. Eleven patients in the Reiter group who had uninterrupted treatment were compared with 9 in the same group for whom 4-, 5-, and 7-day intervals occurred once for each case. The mean gain for the former group was 5.45 and for the latter 23.66; the difference is significant at the 2% level ($t = 2.67$), but is not in the anticipated direction. This difference is not explained by differences in intelligence level, or by a significant inferiority in the "continuous treatment" group immediately after the course of ECT. Although it appears that improvement in the first two weeks following shock was somewhat related to intelligence (the product-moment correlation between gains and Vocabulary score being .40, significant at the 1% level of confidence), differences in Vocabulary and in the preshock Wechsler-Bellevue IQ between patients of the Reiter group having interrupted treatment and those having continuous treatment (.64 and 8.66, respectively) are not significant ($t = .54$ and 1.49). For the memory quotient earned on the final shock day the mean of the "interrupted treatment" group surpassed that of the "continuous treatment" group by only 3.12; the difference is not significant ($t = .55$). We are at a loss to explain why a greater amount of reversible cognitive disturbance should occur to those patients whose treatment, because of one unduly long rest interval, was spread over a longer period of time. There is no evidence that these brief interruptions operated to alleviate the disrupting effect of treatment.

Wechsler-Bellevue Scale

Changes pre- to postshock. Also in Table 2 are shown mean IQs and their standard deviations for two administrations of the Wechsler-Bellevue Scale; Form I was given before the start of ECT and Form II a fortnight after its completion. Though in individual cases the changes may be marked (the range being 30 IQ points lost to 9 IQ points gained), for each group as a whole a slight decrement is found on retest, which is significant, at the 5% level, in the case of the Conventional group and approaches significance for the Reiter group. The experience of Gerboth (3) and of Hayes and Schneider (4) suggests that with normal subjects a negligible gain results when Form II follows Form I. Here the expectation would be that remitting psychosis, present in more than half the cases, would bring about an increase in scores, but, presumably because after two weeks there is still residual impairment of cognitive functions, this effect is not found. Confirming evidence that intellectual loss might still be present on retest is brought out in one qualitative feature, a difficulty with word finding, which occurred in a few instances. In responding to Picture Completion two weeks after the end of the shock series, four patients, though they indicated by pointing that they understood what was missing in item 3, were unable to recall the word *hinge*, and five others recalled it only after a marked hesitation.

Test of Hypothesis III. The difference between the experimental groups in average changes (gains or losses) in Wechsler-Bellevue IQ from first to final administration is negligible. Therefore the hypothesis that pre- to postshock average gains would be greater, and average losses less, for the Reiter than for the Conventional group is not sustained.

Summary

In order to compare the detrimental effect on cognitive functions of two different methods of administering ECT, two groups of 20 patients each, matched for age, education, and intelligence, were subjected to ECT, the first group to 20 convulsions by the Reiter method, and the other to 12 convulsions by the Conventional method. Various psycho-

logical measures were used before shock, during the course of shock, and two weeks postshock.

At the close of the treatment period less cognitive disturbance, as measured by certain tests involving new learning, had occurred in the Reiter group, but differences were negligible for other tasks. In the two-week postshock period gains in the Wechsler Memory Scale were approximately equal for the two groups. Comparison of pre- to postshock average changes in the Wechsler-Bellevue Scale gave no advantage to either group.

Received January 24, 1956.

References

1. Friedman, E. Unidirectional electro-stimulated convulsive therapy. *J. nerv. ment. Dis.*, 1949, **109**, 540–549.
2. Friedman, E., & Wilcox, P. H. Electrostimulated convulsive dazes in intact humans by means of unidirectional currents. *J. nerv. ment. Dis.*, 1942, **90**, 56–63.
3. Gerboth, Renate. A study of two forms of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 1950, **14**, 365–370.
4. Hayes, W., & Schneider, B. A test-retest evaluation of the Wechsler Forms I and II with mental defectives. *J. clin. Psychol.*, 1951, **7**, 140–143.
5. Jellinek, E. M., & Bolles, Marjorie. The effect of bromides on certain psychological test performances. *J. Psychol.*, 1942, **14**, 277–293.
6. Kalinowsky, L. B., & Hoch, P. H. *Shock treatments, psychosurgery, and other somatic procedures in psychiatry*. (2nd Ed.) New York: Grune & Stratton, 1952.
7. *Kuhlmann-Anderson Intelligence Tests*. (5th Ed.) Minneapolis: Educational Test Bureau, 1942.
8. Pascal, G. R., & Zeaman, Jean B. A note on the validity of Wechsler-Bellevue scatter. *Amer. J. Psychiat.*, 1949, **105**, 840–842.
9. Ruesch, J. Intellectual impairment in head injuries. *Amer. J. Psychiat.*, 1943, **100**, 480–496.
10. Stone, C. P. Losses and gains in cognitive functions as related to electro-convulsive shocks. *J. abnorm. soc. Psychol.*, 1947, **42**, 206–214.
11. Stone, C. P., Girdner, J., & Albrecht, Ruth. An alternate form of the Wechsler Memory Scale. *J. Psychol.*, 1946, **22**, 199–206.
12. Summerskill, J., Seeman, W., & Meals, D. W. An evaluation of postelectroshock confusion with the Reiter apparatus. *Amer. J. Psychiat.*, 1952, **108**, 835–838.
13. Wilcox, Katherine W. Confusion and therapy in electroconvulsive treatment. *Conf. Neurol.*, 1954, **14**, 318–326.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Group Psychotherapy with Chronic Psychotic Soiling Patients[1]

John E. Tucker
VA Hospital, Albany, N. Y.

This study was designed to test the effectiveness of group psychotherapy under controlled experimental conditions with an objective criterion. Twenty male psychotic *Ss* without detectable organic bases for their soiling habits were selected from a ward of 38 soilers. A soiler was defined as a psychotic patient with urinary or fecal incontinence.

The frequency of soiling for the 20 *Ss* was measured in a 15-day pre-experimental period using standard tally sheets kept by trained ward nurses. The 20 *Ss* were given a brief test of mental alertness which gave a quantitative score. The *Ss*' ages were taken from medical records. The *Ss* were paired by trial and error into 2 groups of 10 each so that the means and *SD*s of their ages, soiling frequencies, and mental-alertness test scores were not statistically different between the 2 groups (critical ratios for M_d: .590, .614, .870). The *Ss* were diagnosed chronic psychotics. One group was randomly designated the control group and did not have group therapy. The other group was the experimental group and was given 45 hours of group therapy.

Therapy was oriented toward increasing verbal communication. The *Ss* were regressed and disorganized. Transference was fostered by feeding the *Ss* sweets. No overt changes were noted in *Ss*' psychotic behavior but they showed friendliness toward the therapist and did not resist group meetings. The *Ss* achieved some psychological group structure as shown by consistent choice of seats, the formation of submissive-dominant relationships, and some role-taking.

During the 9-week experimental period the 20 *Ss* stayed on their usual ward and did not have any change in their living routine except for group therapy for the 10 experimental *Ss*. The identities of the 10 controls were concealed from ward personnel. The tally of soiling was kept over the 9-week period for both groups. At the end of the period the pre-experimental soiling data were compared with postexperimental and mean gain differences were analyzed. Controls decreased 1.1 in their daily soiling average. Experimentals decreased 6.7 in theirs. The 5.6 difference favoring the experimental group mean gain was statistically significant, p less than .05.

The selection and matching procedures embodied in the design encourage the hypothesis that the greater reduction in soiling behavior in the experimental *Ss* resulted from group therapy. Possibly, group therapy was a sufficiently potent technique to modify early learned behavior patterns such as toilet habits. Use of group therapy with other behavior pattern disturbances found in psychiatric hospitals such as assaultiveness or elopement suggests itself.

Brief Report.
Received July 24, 1956.

[1] An extended report of this study may be obtained without charge from John E. Tucker, VA Hospital, Albany, New York, or for a fee from the American Documentation Institute. Order Document No. 5040 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Short-Term Effects of Chlorpromazine on Psychiatric Patients

William D. Winter and Wilbur K. Frederickson

Hospital for Mental Diseases, Milwaukee, Wisconsin [1]

The increasing use of chlorpromazine and other tranquilizing drugs in the treatment of mental disorders poses practical and theoretical problems for the clinical psychologist working in this area. Experience with these drugs has demonstrated their efficacy in making disturbed patients more manageable and reducing the need for restraints, sedative packs, and seclusion (5). Used in conjunction with psychotherapy, the drugs have replaced to a large extent the more drastic "shock" therapies as the treatment of choice. The exact physiological and psychological action produced by these drugs is still largely unknown. We see clinically a reduction of anxiety, agitation, and emotional lability in the patient which makes him more amenable to psychotherapy. It is still undetermined whether these reactions indicate that basic personality changes are being effected in an individual or whether the drugs only enable the patient to work out his problems more effectively by lessening the confusion and intense emotionality set off by unbearable conflicts.

One purpose of the present study will be a preliminary investigation into the nature of personality changes produced under chlorpromazine, as measured by ward behavior and psychological tests. Our first hypothesis is that within a week's time no significant personality trait changes, as measured by the test used in this study (MMPI), would occur as a direct result of chlorpromazine treatment. We, however, predict an improvement in overt ward behavior, related to a decrease in subjective anxiety.

Another general problem investigated, a practical one for psychologists, concerns the validity of administering tests involving visual-motor coordination to patients undergoing chlorpromazine treatment. In this hospital most patients are routinely put on some form of drug therapy a few days after their admission. Physical symptoms due to the physiological action of the drugs occur shortly after the patients begin this therapy. These symptoms usually include blurred vision, difficulty in coordination, and dizziness. A test involving visual-motor coordination, such as the Bender-Gestalt test, can be expected to be particularly vulnerable to misinterpretation if the side reactions to the drugs significantly affect test performance. Our testing experience with patients undergoing drug therapy led us to predict that tests involving a visual-motor task would be contaminated by the physical side effects of the drug.

Methodology

The subjects used in this study were 59 female patients newly admitted to the Milwaukee County Hospital for Mental Diseases. These patients represented a variety of diagnostic categories, but all those with known or suspected organic impairment or those in need of emergency care were excluded from the sample. The age range of the subjects was from 18 to 45 years, with a median age of 33.5. The median educational level was 11.5 years of schooling, and the median mental age was 15.5, as estimated from the Shipley-Hartford scale.

[1] The authors wish to express their appreciation of the assistance given them in carrying out this study by Christine Ivanov, M.D., Donald Schaefer, M.D., Robert Lichtenstein, M.D., and the nursing staff of this hospital.

The test battery included the Bender-Gestalt test, the group form of the MMPI, and the Shipley-Hartford scale. The Gardner Behavior Chart (8) was used to estimate adjustment to ward routine. A check list showing the deviation of the patient's drawings from the stimulus figures was used to score the Bender-Gestalt test. This check list was developed by the authors at the suggestion of Max Hutt. This list incorporates many of the items in Pascal and Suttell's scoring system (6), and also includes deviations such as constriction and overexpansiveness, which are found in Hutt's criteria (4). Unlike Pascal's system, differential weighting of specific deviations was not attempted. Scoring reliability for this check list was found to be .91. The scoring of the MMPI followed the practice of adding the designated percentage of *K* to the necessary diagnostic scales (3).

As patients were admitted to the hospital, they were randomly assigned to two groups. Both groups received placebos, which were identical to the chlorpromazine tablets in size and color, for three days. At the end of this time, each patient was interviewed and tested with the Bender-Gestalt test and the MMPI, and a nurse who was familiar with the patient filled out the Gardner Rating Chart on her. On the fourth day, the patients in the Drug group began receiving 50 mg. tablets of chlorpromazine (thorazine) 4 times a day, and the Placebo group continued getting placebos. Seven days later, both groups were retested on the Bender and MMPI and another Gardner Rating Chart was filled out. The Shipley-Hartford scale was also administered at this time. At no time did the physicians, nurses, or psychologist doing the testing know which of the two coded treatment routines involved the use of chlorpromazine. The Bender records were also coded so that the scorers would not know whose test was being scored or whether the protocol was from the first or second test administration.

The two groups were not significantly different on the variables of age, education, intelligence, number of admissions to the hospital, or diagnostic category.

Results

In Table 1 are listed the test-retest reliability coefficients for each of our measures with the time interval of one week between tests. These correlations are based on the scores of the drug and placebo groups combined, and are similar to the average reliabilities found in previous studies (1, 3, 7). The median reliability for the placebo group alone was .75; for the drug group, .745.

The changes occurring in the test scores over a one-week period are given in Table 2

Table 1

Test-Retest Reliability Coefficients

Measure	*N*	*r*
Ward Ratings	51	.86
Bender-Gestalt	44	.80
MMPI *T* Scores		
L	31	.78
F	31	.76
K	31	.65
Hs	31	.76
D	31	.79
Hy	31	.66
Pd	31	.71
Mf	31	.63
Pa	31	.75
Pt	31	.86
Sc	31	.77
Ma	31	.65

Table 2

Significance of Changes Occurring in a One-Week Period for Entire Sample

Measure	*N*	Mean score on first testing	Mean change	*p*
Ward Ratings	51	36.8	+3.1	<.01
Bender-Gestalt	44	27.5	−2.6	<.01
MMPI *T* Scores				
L	31	57.2	+1.7	<.20
F	31	60.1	−1.9	<.20
K	31	59.4	+2.5	<.10
Hs	31	58.6	−1.2	<.40
D	31	68.2	−3.2	<.20
Hy	31	65.2	−2.8	<.20
Pd	31	72.0	−4.7	<.02
Mf	31	54.7	−1.5	<.30
Pa	31	59.0	−1.9	<.30
Pt	31	63.0	−3.9	<.02
Sc	31	65.2	−4.1	<.02
Ma	31	55.3	−2.7	<.20

for the entire sample. Change was computed by subtracting the first test score from the second test score of each subject, and the significance of this difference was determined by the use of two-tailed *t* tests for correlated data. As can be seen from the table, our subjects as a group showed significant improvement in ward behavior and Bender-Gestalt scores, and a significant decrease on the Psychopathic Deviate, Psychasthenia, and Schizophrenia scales of the MMPI. They also showed a tendency toward a higher *K* score on the MMPI at the time of their second testing.

Inspection of the data revealed that the two groups differed significantly on several MMPI scales on the original testing, therefore, the drug and placebo patients were matched as closely as possible on the basis of their initial scores on each measure, in order to determine which group changed more in one week's time. Comparison of the mean changes of the two groups by use of the *t* test for uncorrelated means revealed only three variables in which the differences approached significance at between the .10 and .05 level of confidence. The patients receiving drugs improved more than those taking placebos on the Ward Ratings, and the placebo group improved more on the Bender. The drug group showed significantly higher scores on the *K* scale of the MMPI. The number of subjects compared on Ward Ratings was 42, on the Bender-Gestalt 32, and on the *K* scale 22.

Discussion

Analysis of the data indicates that in a period of one week the patients as a group showed small but statistically significant changes in Ward Behavior ratings, Bender-Gestalt scores, and on several MMPI scales. This is congruent with our own clinical experience with acutely disturbed patients, and with the result of other studies (7); the severity of their symptoms often decreases markedly after only a few days of hospitalization. The problem of the depth and permanence of these changes cannot be completely solved. It should be pointed out that our data are limited and we cannot precisely evaluate the effects of practice, the lack of independence of the MMPI scales, the short time interval, and the use of only a moderate drug dosage.

When we compare the mean changes of the drug group with those of the placebo group, we do not find any outstanding differences. There is a trend, however, for the drug group to show better ward adjustment, poorer coordination on the Bender-Gestalt, and a higher *K* score. The first two findings are in the predicted direction. The use of chlorpromazine reduces subjective anxiety and excessive emotional lability, which makes the patient easier to get along with and more approachable. At the same time, particularly in the early phases of drug therapy, the patient suffers from many side reactions such as dizziness, blurred eyesight, poor coordination, and physical unrest, all of which could easily affect performance on the Bender-Gestalt test.

Apparently none of the usual diagnostic scales on the MMPI were differentially affected by the drug. The test-retest reliability figures for the two groups on the MMPI are similar to those cited in the MMPI manual and in other studies (3, 7). This suggests that the use of chlorpromazine therapy with the patients resulted mainly in superficial behavioral improvement which may have involved a suppression of symptoms rather than deep personality changes. There could, of course, be an increase in ego strength which may lead to eventual major changes in personality structure, but the data shed no light on the question. It is interesting that other studies in this area have shown little or no change in psychological tests after drug therapy despite such overt changes in behavior as to markedly lessen the need for sedative packs, electric shock, seclusion, etc. (2, 5).

A comment should be made concerning the need for using control groups when doing research in this area. As mentioned above, a certain percentage of acutely ill patients, even when not given any special kind of treatment except for routine hospitalization, tend to get better. When control groups cannot be used, it is easy to attribute this improvement to whatever therapy is being administered. In one case a physician hesitated

about putting a patient on the project because she was quite disturbed and he wanted her to receive electric shock. She was nevertheless placed on the project where she made spectacular improvement in her ward behavior. The physician was reassured that the patient was at least getting some drug therapy, but a later check of the records showed that the woman was in the placebo group. Thus the physician was falsely attributing her spontaneous (i.e., of unknown cause) recovery to the action of the drug.

In interviews with patients we were made aware of the effect of group suggestion. Many patients entered the hospital after reading optimistic newspaper reports of the "wonder drugs," and every patient was exposed to the enthusiasm of the medical staff and to the comments of other patients. Sometimes a patient in the placebo group would report not only a lessening of tension due to the supposed drug she was receiving, but would also dwell on the physical discomfort the medication caused.

Improvement in both groups varied considerably from individual to individual. For example, the change in ward behavior rating on the placebo group ranged from − 12 to − 15, and for the drug group from − 4 to − 14.

The drug, then, can be said to cause some apparent changes for the better in a seven-day period, but it is not possible to say how central these changes are or how deep they go. Individual differences are great, and the therapeutic effects may sometimes be exaggerated in uncontrolled research by neglecting the factors of spontaneous improvement, suggestion, and biased observation.

Summary

A group of acutely disturbed female patients was tested before and after a week of chlorpromazine therapy, and the changes in test scores were compared with those of a control group which received only placebos. The combined groups showed small but significant changes in Ward Ratings, Bender-Gestalt test scores, and on the *Pt*, *Pd*, *Sc*, and *K* scales of the MMPI. When patients were matched on their original score for each measure, the data showed a trend of the drug group to improve more in Ward Ratings and increase more on *K*, but to do less well on the Bender-Gestalt. Our conclusions are that, in general, acutely disturbed patients tend to improve after a few days of hospitalization and that chlorpromazine speeds behavioral adjustment to ward routine, but that it causes some disruption of coordination, especially during the so-called "turmoil phase" at the beginning of treatment. Evaluation of psychological tests involving visual-motor coordination, such as the Bender-Gestalt, should, therefore, take into account the effects of concurrent drug therapy.

Received April 9, 1956.

References

1. Addington, M. A note on the Pascal and Suttell scoring system of the Bender-Gestalt test. *J. clin. Psychol.*, 1952, **8**, 312–313.
2. Cowden, R. C., Zax, M., & Sproles, J. A. Group psychotherapy in conjunction with a physical treatment. *J. clin. Psychol.*, 1956, **12**, 53–56.
3. Hathaway, S., & McKinley, J. C. *Minnesota Multiphasic Personality Inventory: Manual.* New York: Psychological Corp., 1951.
4. Hutt, M. L. Revised Bender Visual-Motor Gestalt Test. In A. Weider (Ed.), *Contributions toward medical psychology*. New York: Ronald, 1953. Pp. 660–688.
5. Kasak, M., Schaefer, D., & Winter, W. D. The use of chlorpromazine in a psychiatric hospital: one year's experience. *Marquette Med. Rev.*, 1955, **21**, 17–24.
6. Pascal, G. R., & Suttell, Barbara J. *The Bender Gestalt Test; Quantification and validity for adults.* New York: Grune & Stratton, 1951.
7. Rosen, A. Test-retest stability of MMPI scales of a psychiatric population. *J. consult. Psychol.*, 1953, **17**, 217–221.
8. Wilcox, P. H. The Gardner Behavior Chart. *Amer. J. Psychiat.*, 1942, **98**, 874–880.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Use of the Proverbs Test for Differentiating Schizophrenics from Normals[1]

Donald R. Gorham
Veterans Administration Center, Waco, Texas
Baylor University, Waco, Texas

In the Proverbs Test (10), the subject is asked to tell the meaning of a proverb. For normal subjects, even as young as fifth graders, this request readily stimulates what Goldstein refers to as the "abstract attitude." Indeed, the process of translating the concrete symbols of a proverb into generalized concepts is almost explicitly described in one of Goldstein's characterizations of abstraction:

> In the abstract attitude, we transgress the immediately given specific aspect or sense impression; we abstract from particular properties. We detach ourselves from the given impression, and the individual thing represents to us an accidental sample or representative of a category (7, pp. 18–19).

Benjamin (3) believed that the breakdown of verbal abstraction so commonly present in schizophrenia could be investigated by proverbs. Goldstein (8) claimed that the findings of the test for abstraction and concreteness which he and his associates had developed for investigating organic damage had relevance for research with schizophrenia and pointed out the need for more measures of abstraction.

The clinical Proverbs Test consists of three parallel forms containing 12 proverbs each, scaled in the order of difficulty. Samples are:

1. Rome was not built in a day.
2. When the cat's away the mice will play.
3. The wife is the key to the house.
4. Speech is the picture of the mind.
5. Better be happy than wise.
6. The more cost, the more honor.

An abstract (*Ab*) score on this form of the test is secured by rating the items according to sample scores for each proverb.

The multiple-choice Proverbs Test consists of 40 items with four answers each. It is scored by stencil. An abstract (*Ab*) score consists of the total number of correct responses. A concrete (*Co*) score consists of the number of items marked from 18 "concrete" answers. Samples of an easy and a difficult item are:

1. The sun shines upon all alike.
   a. It's the same sun everywhere.
   b. All are created equal. (*Ab* answer)
   c. The sun shines on everybody. (*Co* answer)
   d. People that do the same things are alike.
2. The good is the enemy of the best.
   a. Get the best of your enemy whenever you can.
   b. It is best to be good even if you fail.
   c. We should love our enemies.
   d. Most people are too easily satisfied. (*Ab* answer)

The Proverbs Test seems to possess useful qualities as an instrument for measuring impairment of conceptualization in schizophrenia. It gives a relatively pure measure of "verbal Comprehension" as evidenced by high correlations with tests heavily loaded with this factor (9, p. 7).

These considerations led the author to explore the possibilities of utilizing the Proverbs Test to investigate impairment in abstraction in schizophrenia and to differentiate schizophrenics from normals. Reliability studies indicated that the test was satisfactory when the subjects were hospitalized schizophrenic patients (10). As samples of schizophrenic populations were gathered, it be-

[1] A paper presented at the annual meeting of the Southwestern Psychological Association, Dallas, Texas, March 23, 1956.

came apparent that their Proverbs Test scores were generally lower than those of normals. Since test scores are so closely correlated with intelligence, however, this factor had to be controlled before it could be determined whether an actual difference existed.

A pilot study was made in which 78 normal subjects were compared with 118 chronic schizophrenics. Analysis of covariance was used to hold intelligence constant and the Wechsler vocabulary subtest was used as the measure of intelligence. The analysis yielded an F of 14.30 which is significant at the .001 level of confidence. This finding seemed to justify a rigid test of the variation of scores on the Proverbs Test between schizophrenics and normals.

Method

From a sample of 332 Air Force basic airmen and 232 hospitalized schizophrenic veterans, 100 AF men were matched for sex, education, and intelligence with 100 chronic schizophrenics. The Word Knowledge Test of the Airman Classification Battery was used as the measure of intelligence. The pairs were matched exactly for stanine scores on the Word Knowledge Test. They were matched within two years for level of education; about two thirds matching exactly. The findings of this study are presented in Table 1.

The abstract (*Ab*) score and the concrete (*Co*) score on the multiple-choice Proverbs Test are secured by the use of scoring stencils. Comparable scores for the individual (clinical) form of the test are secured by individual scoring according to scoring samples for each proverb (9, 10). The table shows that the differences in the means of both the abstract and the concrete scores for both forms of the test are significant beyond the .001 level of confidence.

Since intelligence has been held constant by matching each normal with each patient on the Word Knowledge Test, the striking drop from normal to schizophrenic group on the abstract scores and the corresponding rise in the concrete scores indicate an impairment in abstraction resulting from the schizophrenic condition of the patients. In this connection it is interesting to note the negative correlation between the abstract and concrete scores on the Proverbs Test where independently derived. The coefficients of correlation between the *Co* of the multiple-choice test and the *Ab* score of the clinical test were found to be $-.58$ for 100 normal subjects and $-.54$ for 173 chronic schizophrenic subjects.

Cutting Points and Their Application

In order to establish cutting points and norms for use with the Proverbs Test the total available population of normals and schizophrenics was divided into intelligence groupings according to scores on the Word Knowledge Test. Such treatment of the data was necessitated by the correlation between

Table 1

Validity of the Proverbs Test for Discriminating Between Normal and Schizophrenic Men with Education and Intelligence Held Constant

(100 matched pairs: $N = 200$)

Test and score	Air Force enlistees M	Air Force enlistees *SD*	Chronic schizophrenics M	Chronic schizophrenics *SD*	Mean difference D	*t*	*p*
Multiple-Choice Test							
Ab (Abstract) score	24.6	7.4	16.7	8.7	7.9	10.34	$\ll .001$
Co (Concrete) score	1.4	2.3	4.8	3.5	3.4	9.50	$\ll .001$
Clinical Test (Form I)							
Ab (Abstract) score	11.5	6.3	5.2	6.5	6.3	34.45	$\ll .001$
Co (Concrete) score	1.8	2.6	4.7	3.5	2.9	32.22	$\ll .001$

Table 2

A Comparison of Scores of Normals and Schizophrenics
Means and *SD*s by Intelligence Groupings *

Test and score	*Ab* (abstract) score		*N*	*Co* (concrete) score	
	Mean	*SD*		Mean	*SD*
Multiple-choice test					
High Intelligence (W.K.: 7, 8, 9)					
Normals	31.3	3.9	58	0.3	0.7
Schizophrenics	24.8	7.3	65	2.8	2.9
Middle Intelligence (W.K.: 4, 5, 6)					
Normals	22.8	5.4	136	1.6	2.4
Schizophrenics	15.3	6.5	79	5.1	3.6
Low Intelligence (W.K.: 1, 2, 3)					
Normals	16.0	4.9	58	3.9	3.0
Schizophrenics	9.6	3.9	65	6.1	2.9
Clinical test					
High Intelligence (W.K.: 7, 8, 9)					
Normals	19.2	2.6	58	0.2	0.6
Schizophrenics	12.6	6.4	65	1.9	5.9
Middle Intelligence (W.K.: 4, 5, 6)					
Normals	11.4	5.7	136	1.1	1.6
Schizophrenics	4.9	4.7	79	2.5	4.4
Low Intelligence (W.K.: 1, 2, 3)					
Normals	4.0	4.0	58	3.1	2.7
Schizophrenics	1.0	1.0	65	3.9	2.4

Note.—The normal population consisted of Air Force basic airmen from the Personnel and Training Research Center, Lackland A. F. Base. The schizophrenic population consisted of a representative sample of testable chronic hospitalized veterans from Waco Veterans Administration Center. These data are taken from norms in the clinical manual of the Proverbs Test (10) and are used with the permission of the publisher.

* Based on the Word Knowledge Test of the A. F. Airman Classification Battery. The W. K. scores 7–9 include the top 23%, 4–6 include the middle 54% and 1–3 include the bottom 23%.

intellectual level and performance on the Proverbs Test. The norms by intelligence groupings are presented in tables in the clinical manual of the Proverbs Test (10). In these tables the mean score and *SD* of normals and schizophrenics can be compared for each stanine score group of the Word Knowledge Test. Some of these data are presented in Table 2, which shows combined groupings of stanines 1–3, 4–6 and 7–9 on the Word Knowledge Test. These divisions are roughly equivalent to the upper quartile, middle half, and lower quartile in a distribution of normal subjects.

The data in Table 2 were used as the basis for establishing experimental cutting points for differentiating normals from schizophrenics. Actually, a considerable amount of trial and error entered into the process. It was found that at the upper and middle intelligence levels, the abstract and concrete scores of the multiple-choice test could be used most efficiently in combination. This was not true at the low intelligence level, where the concrete score had no usable differentiating value (see Table 2). The effectiveness of the cutting points when applied to a population of 352 normals and 232 chronic schizophrenics is shown in Table 3. The fact that a substantial portion of these cases was included in the population from which the cutting points were derived should be taken into consideration in judging the effectiveness of this application.

Table 3

Application and Cross Validation of Cutting Points for Multiple-Choice Test

Group and cutoff scores	Original sample			Cross validation sample		
	Normal	Schizophrenic	Total	Normal	Schizophrenic	Total
High Intelligence (W.K.: 7–9)						
$Ab \geq 26$ and $Co < 2$	86	22	108	48	6	54
$Ab < 26$ and/or $Co \geq 2$	19	43	62	9	21	30
	105	65	170	57	27	84
Middle Intelligence (W.K.: 4–6)						
$Ab \geq 15$ and $Co < 4$	112	23	135	24	8	32
$Ab < 15$ and/or $Co \geq 4$	29	56	85	4	22	26
	141	79	220	28	30	58
Low Intelligence (W.K.: 1–3)						
$Ab \geq 12$	70	18	88	16	10	26
$Ab < 12$	16	70	86	4	22	26
	86	88	174	20	32	52
Combined Groups						
Predicted Classification						
Normal	288	63	351	88	24	112
Schizophrenic	64	169	233	17	65	82
	352	232	584	105	89	194

The table shows that there were 64 false positives in the normal population sample and 63 false negatives in the schizophrenic population sample. The predictive efficiency of the application of the cutting points was: for normals, 81%, for schizophrenics, 73%, and for the total population of 584 cases, 77.5%.

In attempting to establish cutting points for the clinical test, it was found that the concrete scores could not be used successfully in combination with the abstract scores. The latter scores used alone, however, separated normals from schizophrenics with about the same degree of efficiency as the multiple-choice test scores. The application of the cutting points to the 100 matched pairs described earlier in this paper is shown in Table 4. The table shows that there were 17 false positives in the normal group, 28 false negatives in the schizophrenic group or a total of 45 misclassifications in the total sample of 200 cases. The efficiency of the clinical test for this sample is thus 77½%.

Cross Validation

The cutting points for use with the multiple-choice test were applied to a new population of normals and schizophrenics in order to cross validate the effectiveness of the measures. The new normal population consisted of 105 men and women in civilian and civil service employment at a variety of vocational levels. The new schizophrenic population consisted of 51 male veterans and 38 male and female chronic state hospital patients. The results of the cross-validation application of the cutting points for the multiple-choice test are given in Table 3.

The predictive efficiency of the cutting points applied to this new sample of 194 cases was 78%—almost identical with the original findings. The new sample contained men and women in both groups, confirming the normative data on the Proverbs Test which showed no sex differences (9). The inclusion of state hospital patients in the new sample indicates that the experimental cut-

Table 4

Application of Cutting Points for Individual Test *

Group and cutoff scores	Normal	Schizophrenic	Total
High Intelligence (W.K.: 7–9)			
Abstract score ≥ 14	24	8	32
Abstract score < 14	5	21	26
	29	29	58
Middle Intelligence (W.K.: 4–6)			
Abstract score ≥ 8	38	14	52
Abstract score < 8	10	34	45
	48	48	96
Low Intelligence (W.K.: 1–3)			
Abstract score ≥ 2	21	6	27
Abstract score < 2	2	17	19
	23	23	46
Combined Groups			
Predicted Classification:			
Normal	83	28	111
Schizophrenic	17	72	89
	100	100	200

* Population; 100 normals (Air Force basic airmen); 100 schizophrenics (chronic hospitalized veterans).

ting points have validity when applied to other than the male veteran schizophrenic patients on the basis of whose scores they were evolved.

Discussion

Several questions are raised in connection with the utilization of the Proverbs Test for differentiating schizophrenics from normals. The comparatively poor results of the multiple-choice test at the high intelligence level stand out. An examination of the actual cases in the schizophrenic group showed that 56% of the false negatives were paranoid schizophrenics, whereas only 34% of the true negatives fell in this diagnostic group. This finding points to the relative ineffectiveness of the multiple-choice test for screening out paranoid schizophrenics.

It is recognized, of course, that the efficiency of prediction is affected by the type of population to which the method is applied. A population containing roughly 50% normals and 50% schizophrenics does not actually exist. If applied to an unselected population containing only 1% schizophrenics, the predictive efficiency would be greatly reduced. On the other hand, if applied to a clinic population consisting entirely of people seeking assistance for mental and emotional problems, the predictive efficiency might well be enhanced.

When used with schizophrenics, the free-answer clinical test differs from the multiple-choice test in that it elicits psychotic material. In addition to concreteness in thinking which is associated with a breakdown of the abstracting function, the schizophrenic frequently "escapes" by a further regression characterized by "originality of performance" (1, p. 317). This is evidenced on the clinical test by responses characterized by autism, bizarreness, irrelevance, perseveration, and delusional preoccupation. Because of this free flow of schizophrenic material which is stimulated by this form of the Proverbs Test, diagnosis of schizophrenia is sometimes obvious upon clinical evaluation even though the objective scores fail to properly classify the subject.

It should be re-emphasized that the Word

Knowledge Test, which was used as the measure of intelligence in this investigation, is highly correlated with the Proverbs Test and gives a relatively pure measure of "verbal comprehension" (9, Table 3). Other testers will find the basic data for establishing intelligence divisions and cutting points suitable for their own purposes in the clinical manual of the Proverbs Test (10).

Summary and Conclusions

1. In a rigid statistical test to determine its effectiveness for discriminating between normal and schizophrenic subjects, both multiple-choice and individual clinical forms of the Proverbs Test yield mean scores whose differences are highly significant.

2. Cutting points for differentiating normals from schizophrenics are presented for three intelligence groupings which approximate the top quartile, middle half, and lowest quartile. The application of these cutting points correctly screens out approximately 80% of normal subjects and 75% of schizophrenic subjects. This level of efficiency was maintained on a cross-validation study based on 194 cases.

3. The effectiveness of the screening is consistent in the three intelligence levels for normal subjects. For schizophrenics the screening is least effective for the high intelligence grouping, probably due to the ineffectiveness of the technique when applied to chronic paranoid schizophrenics.

Received April 3, 1956.

References

1. Arieti, S. *Interpretation of schizophrenia.* New York: Brunner, 1955.
2. Baumgarten, F. A proverb for attitude measurement. *Personnel Psychol.*, 1952, **5**, 249–261.
3. Benjamin, J. D. A method for distinguishing and evaluating formal thinking disorders in schizophrenia. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia.* Berkeley: Univer. of California Press, 1944. Pp. 65–90.
4. Bolles, M., & Goldstein, K. A study of the impairment of "abstract behavior" in schizophrenic patients. *Psychiat. Quart.*, 1938, **12**, 42–65.
5. Cohen, J. A comparative analysis of the factors underlying intelligence test performance of different neuropsychiatric groups. Unpublished doctor's dissertation, New York Univer., 1950. Univer. Microfilms, 1950, No. 1923.
6. Fruchter, B. The factorial content of the Airman Classification Battery. *USAF, Personnel Train. Res. Cent., Res. Bull.*, 1949, No. 49–1.
7. Goldstein, K. Methodological approach to the study of schizophrenic thought disorder. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia.* Berkeley: Univer. of California Press, 1944. Pp. 17–40.
8. Goldstein, K. The significance of special mental tests for diagnosis and prognosis in schizophrenia. *Amer. J. Psychiat.*, 1939, **96**, 575–588.
9. Gorham, D. R. A proverbs test for clinical and experimental use. *Psychol. Rep. Monogr. Suppl.*, 1956, No. 1.
10. Gorham, D. R. *Proverbs Test. Clinical Forms I, II and III. Best Answer Form. Manual. Clinical Manual.* Louisville, Ky.: Psychological Test Specialists, 1956.
11. Hanfmann, E., & Kasanin, J. Conceptual thinking in schizophrenia. *Nerv. ment. Dis. Monogr.*, 1942, No. 67.
12. Hayakawa, S. I. *Language in action.* New York: Harcourt Brace, 1941.
13. King, W. R. Ability to abstract. In F. A. Mettler (Ed.), *Selective partial ablation of the frontal cortex: a correlative study of its effects on human psychotic subjects.* New York: Hoeber, 1949. Pp. 218–238.
14. Loeb, E. The function of proverbs in the intellectual development of primitive peoples. *Scientific Mon.*, 1942, **74**, 100–104.
15. Meadow, A., Greenblatt, M., & Solomon, H. C. "Looseness of association" and impairment in abstraction in schizophrenia. *J. nerv. ment. Dis.*, 1953, **118**, 27–35.
16. Pounders, C. J. A study of the Proverbs Test as a measure of the abstract level of concept formation. Unpublished master's thesis, Baylor Univer., 1956.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Ability of Schizophrenics to Modify Responses in an Interpersonal Situation[1]

M. David Diamond
Riverside Hospital, The Bronx, New York

The present paper discusses the results of an experiment designed to test the hypothesis that schizophrenics, as compared to nonschizophrenic emotionally disturbed individuals, are deficient in the ability to modify their responses in the direction of the responses of another person.

Mead defined rationality as the process whereby the individual "can take the attitudes of the others and control his action by these attitudes, and control their action through his own" (2, p. 339). Thus the "rational" person can modify his responses in the light of the responses of others.

Along similar lines of thought, Cameron states that, for the normal person:

> . . . the more effectively he is able to allow the attitudes and responses of others, which he predicts in symbolic role-taking, to influence his own reactions, the more competent he ought to be in social situations (1, p. 93).

Specifically the hypothesis tested was that schizophrenic individuals modify their responses less in accordance with the responses of another person when these responses are communicated in an interpersonal situation than do nonschizophrenic individuals. This hypothesis was tested at two levels of remoteness from immediate stimulation by having the subjects respond, first in the presence of another individual, and subsequently alone.

If an individual in an ambiguous situation modifies his responses in accordance with the responses of another individual, when he is in the presence of the other person, this would indicate that he can at least momentarily share the points of view of another person. If the individual modifies his responses in accordance with the responses of another individual, when the other individual is no longer physically present, this would indicate that his ability to accept the other person's points of view is not merely a transitory one, and that he is able to internalize the points of view of another. As a rough analogy we might, from a psychoanalytic point of view, consider the first task to be a measure of ego controls and functioning and the second one of superego controls and functioning.

Sherif's work with the autokinetic phenomenon suggested a technique for demonstrating the degree to which one's responses are modified by the responses of another, and the degree to which the responses of the other person are internalized. The autokinetic phenomenon consists of the perception of apparent movement of a stationary point of light, and is measured by having the subjects estimate the distance through which the point of light seems to move. Sherif found that when normal individuals are faced with the ambiguous autokinetic phenomenon, they form their own frame of reference and observe the point of light to move within a certain range. When individuals first face the autokinetic situation alone and then face it again in a group situation, they tend to change their estimates toward a norm and range which is established in the group. When subjects first face the autokinetic situation as members of a group, a norm and range are established which are peculiar to the group, and after-

[1] This paper is a condensation of a part of a doctoral dissertation completed at New York University. The author wishes to acknowledge the help of Drs. L. Crafts, D. Wechsler, M. Deutsch, I. Katz, and P. Stout.

wards when the subjects face the same situation alone, they tend to adhere to the norm and range of the group. Sherif views the entire process as a prototype of the formation of a social norm, which then acts as a frame of reference for determining later reactions (3).

The autokinetic phenomenon seemed to be particularly appropriate for studying the effect of social interaction on behavior because it enables the investigator to isolate the variable of social influence from complex field influences such as the effects of learning and experience. Thus, in the present study an autokinetic experiment was used as a measure of the degree to which subjects modified their responses in the direction of the responses of others.

Subjects

The experimental subjects were thirty hospitalized male schizophrenics. Of the thirty schizophrenic subjects ten had been classified as schizophrenic reaction, paranoid type; nine as schizophrenic reaction, childhood type; seven as schizophrenic reaction, simple type; and four as schizophrenic reaction, chronic undifferentiated type.

A control group was sought which manifested a variety of types of psychopathology in order to eliminate the possibility that the hypothesized deficiencies in interpersonal skills were concomitant with emotional disturbance per se. Male drug addicts in a rehabilitation hospital were selected as control subjects since, as a group, they have been found to present a variety of emotional disturbances. No schizophrenic drug addicts were selected. Of the thirty drug addicts, eight were classified as inadequate personality, eight as schizoid personality; six as passive-aggressive personality; and one was classified in each of the following eight categories: adjustment reaction of adolescence, adjustment reaction of childhood, mixed psychoneurosis, neurotic character disorder, personality pattern disturbance, personality trait disturbance, psychoneurotic reaction, and sociopathic personality.

Table 1

Nonparametric Median Test for Initial Differences Between Pair-Members' Estimates of Light Movement When Judging Alone

Subjects	Below median	Above median	Total
Schizophrenic pairs	7	8	15
Drug addict pairs	8	7	15
Total	15	15	

Chi square = .13 $df = 1$ $p > .10$

Table 2

Nonparametric Median Test for Differences Between Pair-Members' Estimates of Light Movement When Judging Together

Subjects	Below median	Above median	Total
Schizophrenic pairs	4.5	10.5	15
Drug addict pairs	10.5	4.5	15
Total	15	15	

Chi square = 4.8 $df = 1$ $p < .02$

The control subjects were selected to be similar to the schizophrenics with respect to age, highest school grade completed, intelligence, race and socioeconomic background. The mean ages of the schizophrenic and drug addict groups were 19.3 and 18.5, respectively; the mean highest school grade completed, 9.6 and 9.9, respectively; the mean IQs 94.1 and 98.8, respectively. The two groups did not differ significantly with respect to any of these three characteristics. Both groups consisted of about 50 per cent Caucasian, and 25 per cent each of Puerto Rican and Negro individuals, primarily of lower socioeconomic background.

Procedure

The modification of responses in accordance with the responses of others was measured by means of an autokinetic experiment, in which the subject judged the apparent amount of movement of a stationary point of light in a darkened room, first when alone, then with another individual, and then finally alone again. Each subject made twenty estimates when alone. On the basis of these initial estimates when judging alone, subjects in

each group were then assigned to pairs so that the initial differences between pair-members' estimates were approximately the same for subjects in both groups. The subject then made twenty judgments when with his pair-member and finally twenty more judgments when alone again.

The relative amount of convergence of pair-members' estimates when judging together provides a measure of the degree of modification of responses when in an interpersonal situation. The relative amount of convergence of estimates of pair-members when later judging alone provides a measure of the degree to which the "group norm" was introjected or internalized.

For each of the three parts of the experiment—individual judgments, judgments in pairs and subsequent individual judgments—the median of the twenty estimates was determined for each subject. The difference between the medians of pair-members' estimates for subjects in the two groups was then compared by means of the nonparametric median test for each part of the experiment. The nonparametric median test is essentially a chi-square test.

Results

Table 1 presents the results of the nonparametric median test which tests the hypothesis that the initial difference between pair-members is the same for the two groups. The resulting chi-square value of 0.13, which is not statistically significant, supports the null hypothesis that the two groups come from populations having the same median difference between pair members with respect to their assignment to pairs. Thus the pairs in the two groups can be considered similar with respect to initial differences between the medians of their estimates, before the experimental variable was introduced.

When judging with the pair members, subjects in both groups modified their responses in accordance with those of the pair members. However, the modifications made by the drug addicts were greater than those made by the schizophrenics.

Table 2 presents the results of the median test for the differences between median estimates of pair members when judging together. The obtained chi-square value of 4.8 is significant at the .02 level. Thus, the data support the hypothesis that schizophrenic individuals modify their responses in accordance with the responses of another person less than do nonschizophrenic drug addicts, when responding in the presence of the other person.

The results obtained with the drug addicts are similar to those obtained by Sherif when he performed an autokinetic experiment with pairs of normal subjects. When responding in a group after having first responded alone, the median difference between Sherif's pair-members' estimates was about 0.50 of an inch, which is the value obtained for the drug addict pairs (3).

When responding alone after having judged with the pair member, the drug addicts continued to modify their responses in accordance with those of the pair member as much as when they were in the presence of the pair member. The schizophrenics, on the other hand, tended to deviate more from their pair-members' estimates when no longer in the immediate presence of the pair member.

Table 3 presents the results of the nonparametric median test for the median difference between pair-members' estimates when judging alone after judging in pairs. The obtained chi-square value of 6.5, which is significant at the .01 level, supports the hypothesis that, after having participated in an actual interaction with another person, schizophrenic individuals modify their responses less in accordance with the responses of the other person than do nonschizophrenic drug addicts.

Table 3

Nonparametric Median Test for Differences Between Pair-Members' Estimates of Light Movement When Judging Alone After Judging Together

Subjects	Below median	Above median	Total
Schizophrenic pairs	4	11	15
Drug addict pairs	11	4	15
Total	15	15	

Chi square = 6.5 $df = 1$ $p < .01$

In this part of the experiment, also, the drug addict pairs resemble Sherif's normal pairs (3), in that the median difference between pair-members' estimates when judging alone after judging together was about 0.50 of an inch for both Sherif's group and the drug-addict group.

Discussion of Results

The results support the hypothesis that schizophrenic individuals modify their responses less in accordance with the responses of another person, when these responses are communicated in an interpersonal situation, than do nonschizophrenic individuals.

The subject's estimates of light movement when judging with another person, might be considered, from a psychoanalytic point of view, to reflect the adequacy of the subject's ego functioning. The person whose ego functions adequately would be expected, in the ambiguous autokinetic situation, to utilize the only frame of reference available besides his own—the responses of the other person—as a guide in making his own subsequent estimates. The results indicate that the schizophrenic has difficulty in utilizing such an external frame of reference.

The experimental evidence would suggest, further, that the schizophrenic does not adequately internalize the points of view of others. Since the superego is that aspect of the personality that is largely developed through a continuous process of introjection, these results suggest that, in the schizophrenic, not only ego functioning, but also superego functioning is affected.

Summary

A variation of the autokinetic experiment was employed to test the hypothesis that, as compared to nonschizophrenic emotionally disturbed individuals, schizophrenics are deficient in the ability to modify their responses in the direction of the responses of another person. Thirty hospitalized male schizophrenics and thirty male drug addicts judged the apparent amount of movement of a stationary point of light in a darkened room, first when alone, then with another individual from his group, and then when alone again.

The results supported the hypothesis. When estimating light movement in the presence of another person, the convergence of the schizophrenic pairs was significantly less than that of the nonschizophrenic drug-addict pairs. When responding alone, after having responded in the presence of the other person, the convergence of the schizophrenic pairs again was significantly less than that of the nonschizophrenic drug addict pairs.

Received February 21, 1956.

References

1. Cameron, N. *The psychology of behavior disorders.* Boston: Houghton Mifflin, 1947.
2. Mead, G. H. *Mind, self and society.* Chicago: Univer. of Chicago Press, 1934.
3. Sherif, M. A study of some social factors in perception. *Arch. Psychol. N. Y.*, 1935, No. 187.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Concept Learning in the Educationally Retarded Child of Normal Intelligence

Nissim M. Levy and Joseph M. Cuddy
Brown University

The report presented here is part of a larger project concerned with the development of a technique for predicting which children of normal intelligence, as measured by the usual methods, are likely to develop learning difficulties. The commonly-used measures in this area are estimates of intelligence, reading readiness, and social and physical development. Unfortunately these measures, except when they are fairly low, are relatively poor predictors.

Since most of the important tasks that children must master in school involve learning, a good instrument for estimating rate of learning may be of considerable value for evaluation purposes. This approach to the problem may be especially fruitful since rate of learning is very likely a complex function of not only intelligence, but also strength of motivation to learn, and reactions to failures during the acquisition process.

A careful examination of typical intelligence tests leads one to the fact that most of the material depends heavily upon the achievements of the individual. Of the WISC subtests, for example, none of them with the possible exception of the coding allow the subject an opportunity to improve performance during the assessment situation. This behavior sample is probably too small to be adequately used as a learning measure. However, it is interesting to note that Altus (1) reported consistently low scores on this subtest in records of 24 severe cases of educational retardation.

The present study was designed to answer the following question: Given a group of children of normal intelligence some of whom are underachievers in school, do underachievers differ significantly from normal achievers in ability to solve a concept-formation problem? A concept-formation problem was chosen for the first investigation on the assumption that it might provide the most sensitive measure of impairment in learning attributable to motivational and emotional difficulties.

Method and Procedure

The *Ss* were 23 pairs of fourth-grade elementary school children matched for age, sex, and, roughly, for socioeconomic status of parents.[1] The *Ss*' IQs ranged from 98 to 103 as measured by the California Test of Mental Maturity. One student of each pair was working up to grade placement while the other was behind from .5 to 2.5 years in achievement as measured by the Stanford Achievement Battery.

The learning task used for the study was a variation of the Oddity problem. Ten sets of Vigotsky Blocks were used. Each set consisted of three wooden blocks. In each set, the correct block was the odd one in terms of one of three possible dimensions. The three dimensions used were size, color, and shape.

Each *S* was told that his problem was to find out how to tell which was the correct block each time he was presented with a different set. If *S* chose the correct block, he received a chip. If he chose an incorrect block, *E* said, "That's not right," and the set of blocks was removed and another set presented. At the end of the game he received a penny for each five chips he had won.

[1] The authors wish to thank Dr. James Hanley, Superintendent of the Providence City Schools, and Miss Mary Lauth for their cooperation and assistance in making this study possible.

Five blocks of ten trials were given to each *S*, each block consisting of a random order of the ten different sets. If *S* chose correctly on 11 consecutive trials, testing was discontinued and *S* was credited with correct choices for the balance of the 50 trials. Testing was continued even when *S* spontaneously verbalized the correct principle before the end of the 50 trials. At the end, all *Ss* were asked to verbalize how they solved the problem.

The measures of learning used were: (*a*) number of errors; (*b*) number of trials to criterion; and (*c*) verbalization of the Oddity principle.

Results

Errors. In 18 out of the 23 pairs of *Ss* in the study, the normal achiever made less errors than his or her matched underachiever. Unless otherwise indicated all results reported were tested for significance by a two-tailed application of the binomial expansion. The probability of getting 18 out of 23 differences in one direction is less than the .01 level.

Trials to criterion. Eleven *Ss* reached the criterion of 11 consecutively correct choices within 50 trials. Of these 11 *Ss*, 10 were normal achievers while only one was an underachiever. This result yields a p value of less than .01.

Correct verbalization of the "Oddity" principle. Sixteen *Ss* correctly verbalized the principle. Of these 16 *Ss*, 14 were normal achievers while two were underachievers. This result yields a p value of less than .01.

Ten *Ss* met both criteria. All 10 of these *Ss* were normal achievers. Again this result is significant at well below the .01 level of confidence.

Using either criterion, 13 of the 23 pairs of *Ss* had a difference in favor of the normal achiever, while no difference occurred in 10 pairs. Inspection of the data suggested the presence of a relationship between degree of discrepancy in achievement between the normal achiever and the underachiever of a given pair, and the likelihood of a difference occurring in either criterion in favor of the normal achiever. Therefore the difference score in achievement level for all of the pairs of *Ss* was calculated. The over-all median of the 23 discrepancy scores was calculated and the data analyzed by means of the nonparametric median test. The contingency table arising from this analysis is presented in Table 1. Fisher's exact method for testing the significance of contingency tables with small frequencies was used and resulted in a p value of .01, confirming the suggestion of a relationship between the degree of underachievement and the likelihood of a difference in favor of the normal achiever.

It should be pointed out that two-tailed tests are conservative estimates of significance in this study since the implicit hypothesis being tested calls for a prediction of a directional difference. We would predict from both an empirical and theoretical basis that the underachievers would do less well.

Table 1

Median Test for Discrepancy Scores of Matched Subjects

Condition	Above median	Median or below
Normal achiever reached criterion; underachiever did not.	10	3
Neither *S* reached criterion.	2	8

Exact p value, .012

Discussion

The results of the study are quite clear. Every measure of learning used demonstrated a significant difference in performance between our two groups. The finding that pairs of *Ss* with higher discrepancy scores more often yielded a difference in criterion measures is further support for the hypothesis that underachievers learn more slowly.

Since 29 of the 46 *Ss* failed to meet either criterion it is quite likely that the limit of 50 trials was too low. An examination of the verbalization of the *Ss* who did not solve the problem indicated that many of them were at least on the way toward solution. When asked how they tried to solve the problem, some of them indicated partial solution by

suggesting that it was the smaller one or larger one or the one that was a different shape. The *Ss* were credited with the correct solution only if they indicated all three dimensions or simply said that the correct block was always different from the other two.

A difference in the rate of learning of a concept has been demonstrated in this study. The question arises as to whether this difference is restricted to the conceptual area or whether it may be a more generalized deficit. In an attempt to answer this question the same children are being used as *Ss* in a number of other types of learning problems such as pursuit-rotor, stylus maze, paired associates, mirror drawing, discrimination, and problem boxes.

Whether we find underachievers demonstrating a general deficit in learning rate or whether it turns out to be primarily related to the verbal or conceptual areas the question will still remain whether or not we can detect these potential underachievers in preschool or the first grade.

Much of the thinking of clinicians concerned with the problem of the underachiever suggests that in many cases the familial interpersonal constellation and early history of the child play a considerable role in development of this condition. It would be reasonable to assume then that these children have, by the time they reach school age, already developed those patterns of behavior which will result in underachievement. This suggests that we can hope to find a learning-rate difference in children of normal intelligence by the time they reach school age which may be predictive of later underachievement in school.

It must be emphasized that what we are suggesting is not the use of learning tasks as a substitute for the standard intelligence scales. Hodges (4) following a lead of Hunt's (5) reported a study in which he estimated mental age from a variety of measures of performance in a double-alternation problem. The correlations which Hodges obtained varied between .22 and .74. It should be clear from his results that these two techniques are different.

Binet, it will be recalled, defined intelligence as the ability to profit from experience. Most intelligence tests use the same assumption that Binet used in order to justify themselves. This assumption is that all subjects measured by the particular instrument had the same opportunity for the experiences which are relevant to the material being used as the instrument. That this assumption is a weak one has adequately been pointed out before (2, 3). In part it is the weakness of this assumption which makes the use of such tests relatively inadequate for the prediction of academic success. Perhaps if Binet had worked on his problem after Thorndike, the tests which he would have developed might have been more along the lines of the work we are suggesting.

Summary

Twenty-three pairs of fourth-grade children matched for age, sex, and very roughly for socioeconomic status served as *Ss* in a concept-learning task. The *Ss*' IQs ranged from 98 to 103 as measured by the California Test of Mental Maturity. One *S* of each pair was working up to grade placement while the other was behind from .5 to 2.5 years in achievement as measured by the Stanford Achievement battery. All *Ss* were given a maximum of 50 trials within which to reach the criterion of 11 consecutively-correct choices. A variation of the "Oddity" problem was used. At the end of testing, all *Ss* were asked to verbalize how they attempted to solve the problem. The measures of learning were: (*a*) number of errors; (*b*) number of trials to criterion; and (*c*) correct verbalization of the oddity principle.

In 18 out of 23 pairs of *Ss*, the normal achiever made fewer errors than his matched underachiever. Eleven *Ss* reached the criterion of 11 consecutively correct choices. Sixteen *Ss* correctly verbalized the principle. Fourteen were normal achievers while only two were underachievers. Ten *Ss* met both criteria and all 10 of these *Ss* were normal achievers. All of these results were statistically significant below the .01 level of confidence.

Implications of these results were discussed and the direction of further needed research was suggested.

Received July 26, 1956.
Early Publication.

References

1. Altus, G. T. A WISC profile for retarded readers. *J. consult. Psychol.*, 1956, **20**, 153–154.
2. Eells, K. Some implications for school practice of the Chicago studies of cultural bias in intelligence tests. *Harvard educ. Rev.*, 1953, **23**, 384–397.
3. Davis, A. Social-class influence upon learning. (The Inglis Lecture 1948) Cambridge, Mass.: Harvard Univer. Press, 1948.
4. Hodges, A. Double alternation: A measure of intelligence. *J. consult. Psychol.*, 1956, **20**, 59–62.
5. Nichols, P. C., & Hunt, J. McV. A case of partial bilateral frontal lobectomy. *Amer. J. Psychiat.*, 1940, **96**, 1063–1083.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The TAT and Antisocial Behavior[1]

Kenneth Purcell [2]

Mental Hygiene Consultation Service, Fort Dix, New Jersey

Antisocial behavior may be viewed in terms of disequilibrium between an impulse system and a control system, the latter involving assessments of reality consequences as well as internalized standards serving as restrictions upon behavior. The present study is designed to examine certain disturbances of the impulse-control balance as they are manifested in TAT productions. Lindzey has referred to the "determination of the conditions under which inferences based on the projective material directly relate to overt behavior and the conditions for the reverse" (2, p. 18) as one of the major interpretive problems for the TAT. Studies examining the relationship between fantasy and overt behavior have yielded varied results (5, 9). Neither Murray (5) nor Sanford *et al.* (9) found any substantial direct relationship between the intensity of fantasy aggression and its overt expression. However, in explaining their respective results, each hypothesizes a factor of cultural prohibition serving to prevent the overt gratification of antisocial needs. In order to circumvent the complicating factor of middle-class inhibitions regarding aggressive behavior, Mussen and Naylor (6) studied a group of lower-class boys whose culture presumably encourages rather than inhibits overt aggression. They found strong support for a positive relationship between fantasy aggressive needs and overt aggressive behavior. Moreover, they found suggestive evidence that the amount of punishment anticipated relative to aggressive needs was an important determinant of overt aggression. With the exception of Mussen and Naylor's work, previous investigations attempting to correlate covert needs and overt behavior have tended to neglect the role of inhibiting or defensive forces in modifying the behavioral resultant of covert needs. Mussen and Naylor, however, made no effort to distinguish between external and internal controlling forces.

[1] The author wishes to express his appreciation to Richard Besnia, Carl Eisdorfer, Marvin Greenstein, Maurice Kouguell, and Dr. Philip Kubzansky who collected the data for this study and assisted in its statistical evaluation.

[2] Now at the University of Kentucky.

Hypotheses

The major hypotheses of the present study may be stated with reference to different facets of the impulse-control balance:

A. Impulse System

1. There is a direct relationship between degree of aggressive, antisocial behavior, and the number of aggressive thoughts manifested in TAT stories.

2. The form as well as the number of TAT aggressions is related to the extent of overt aggressive behavior. More aggressive individuals are likely to express TAT aggressions crudely and directly, relatively unhampered by repressive processes.

B. External Control System

3. Overt antisocial behavior varies inversely with the anticipated amount of fantasy punishment from an external source.

4. Among antisocial individuals themes of external punishment are more likely to serve as justifications for aggression than among nonantisocial individuals. This hypothesis is derived from Redl and Wineman's (8) detailed account of the "delinquent ego." By this is meant the ego whose every effort is bent toward obtaining guilt-free and anxiety-free enjoyment of delinquent impulsivity.

C. Internal Control System

5. Overt antisocial behavior varies inversely with the anticipated amount of fantasy punishment from an internal source.

6. Of the two types of punishment, internal and external, anticipated internal punishment is the more significant inhibitor of antisocial behavior.

Method

Fifty-seven male Army trainees between the ages of 17 and 26 with a mean age of 20 were used as subjects (*Ss*) for this experiment. All were psychiatric referrals to the Mental Hygiene Clinic for whom psychological testing was requested. The only patients excluded were those suffering from organic brain damage or those diagnosed as psychotic.

Each *S* was seen individually by one of three clinical psychology specialists who administered cards 1, 3BM, 4, 6BM, 7BM, 8BM, 12M, 13MF, 14, and 18BM of the TAT. A fourth clinical psychology specialist and the author, neither of whom had administered any of the TATs, independently scored the TAT protocols for measures of aggressive impulses with respect to both quantity and quality, anticipated punishment from external agencies, and anticipated punishment from within.

The criteria used for identifying fantasy aggressions and punishments were based on those described by Mussen and Naylor (6) with certain modifications. A fantasy aggression score was derived for each *S* by counting the frequency of instances of all aggression such as the following: fighting, criminal assault, getting angry, criticizing, running away, resisting coercion, being negativistic, lying, cheating, stealing, dominating or restraining someone, rejecting someone. The occurrence of death, illness, or accident to parents or other loved objects in a story was also regarded as fantasy aggression. Since the present study is concerned only with extrapunitive aggression, such items as suicide, self-injury, and self-depreciation were not scored as fantasy aggression. The Pearson *r*, used as the reliability measure in every case, showed the correlation between independent scorings to be .91.

An external punishment score was obtained by summing the frequency of such themes as the following when they were directed toward the hero: assault, injury, threat, quarreling, deprivation of some privilege, object or comfort, domination, physical handicap such as blindness, etc., rejection. Scoring reliability was .81.

Internally based punishment included suicide, self-depreciation, and feelings of guilt, shame, or remorse. Since a broad definition of punishment encompasses "injury to a loved object," instances of death, illness, or accident to parents or other loved objects were recorded in this category of internal punishment. In instances in which the hero of the story was the direct agent of aggression, e.g., husband striking his wife, internal punishment was not scored except where there was a clear expression of remorse or shame. By contrast, accidental injury, death, or illness of a loved one always received a scoring of internal punishment unless the hero specifically denied anything resembling guilt feelings. Scoring reliability was .89.

Remoteness of fantasy aggression from overt behavior was rated on a seven-point scale with one representing the most direct form of fantasy expression. For each instance of fantasy aggression counted in the TAT story, a remoteness rating was made. The following considerations were kept in mind when evaluating remoteness: (*a*) object of aggression (person, society, animal, inanimate object), (*b*) time (present, past, future), (*c*) place (customary habitat, other countries, other planets), (*d*) level (behavior, wish, memory, daydream, nightdream, special states of consciousness such as intoxication, dissociation, drug addiction, insanity), (*e*) social context and instigator (hero instigates aggression, merely tags along with crowd, or even has nothing directly to do with aggression, e.g. death of someone due to disease or accident). The reliability for the number of one and two remoteness ratings scored per *S* was .96.

Measurement of aggressive antisocial behavior was based on the social history data routinely obtained by the psychiatric social work specialists. Such information included a record of *S*'s military delinquencies, e.g. AWOL, refusal to obey orders, assault, etc., as well as *S*'s report of any delinquent civilian behavior, e.g., truancies, gang fighting, stealing, peddling of drugs, etc. The military information was, of course, easily verifiable but the data pertaining to premilitary activities were not, except in certain instances in which records were requested from reformatories or other institutions mentioned by *S*. Several factors, however, indirectly testified to the substantial accuracy of *Ss*' accounts of their civilian behavior. First, the civilian and military behavior data were almost invariably consistent with respect to delinquent characteristics. Nevertheless, it was important to use the premilitary history to get a long-term picture of *S*'s characteristic aggressive pattern, particularly for those individuals whose

time in the military was relatively brief. Second, a subsequent check of psychiatric diagnosis revealed no case in which there was gross discrepancy between the clinical impression of the psychiatrist and *S*'s report of his aggressive behavior pattern. Third, on every occasion in which independent confirmation of civilian behavior was available (15 cases), *S*'s own account was corroborated.

Data pertinent to the variable of aggressive behavior were extracted from each *S*'s social history by the three clinical psychology specialists who had been responsible for testing. The author and one of these specialists rated the behavior extracts for aggressivity on a seven-point scale with one representing the least degree of aggressive behavior. The possibility of knowledge of test scores influencing behavior ratings or vice versa was negligible for two reasons. First, the specialist who served as a judge to check the reliability of the behavior ratings had administered only 25 per cent of the total number of TATs, and second, the behavior ratings were accomplished two months prior to the TAT scoring. Sample extracts of the extremes of the scale are presented below.

Rating of 1: Steady school attendance record through fifth term high school. After leaving school worked his way up to level of assistant pressman in print shop at $110 per week. He was drafted from a firm which had employed him since 1948. Married and living with wife and child for three years. No military or civilian delinquencies reported.

Rating of 7: Long history of truancies, delinquency and classroom misbehavior. In reform school four years. Prior to induction was confined two years eight months for armed robbery. While AWOL came in possession of drugs and jailed one month. Assaulted several people. Enlisted in Army on advice of parole officer to get reduction of sentence. History of eight AWOLs and four courts martial. Now facing a general court martial, five years in prison, and a dishonorable discharge because he wrecked stockade dispensary.

The reliability of these behavior ratings was .91. In order to assure large enough groups for comparison the seven-point scale was then consolidated into a three-point scale, combining ratings of one and two; three, four and five; and six and seven. Hereafter, Group I of 24 *S*s refers to the least antisocial group, Group II of 19 *S*s to the intermediate, and Group III of 14 *S*s to the most antisocial group.

Table 1

Product-Moment Intercorrelation of TAT Variables

	FA	Remoteness FA	Int. Pun. FA	Ext. Pun. FA
Remoteness / FA	.53*			
Int. Pun. / FA	−.72	−.58		
Ext. Pun. / FA	−.45	−.13	.38	
Int. Pun. / Ext. Pun.	−.65	−.13	.89	.01

* An r of .36 is required for significance at the 1% level.

Results

Since the interpretation of results depends in some degree upon the relative independence of the five TAT variables their intercorrelations are presented in Table 1. Most of the relationships are statistically significant as seems logically required by the fact that complementary aspects of the impulse-control factor of personality are represented. However, in nearly all instances, the amount of variance in one variable that can be accounted for by the variance in another (as measured by r^2) is very low to moderate. Therefore, each of the variables may be said to contribute some unique information beyond their common meaning.

As shown in Table 2 the analysis of variance indicates over-all significant differences between the means of the three groups. This does not tell us, however, how well the three groups are separated. Therefore, before describing these results in detail, more precise information on the differences between the groups is presented in Table 3. Two-tailed tests of significance were used. It is apparent that in nearly all instances each group is significantly different from the other two.

The difference between the means of the three groups in number of fantasy aggression themes, as shown in Block A of Table 2, is highly significant beyond the value of $p =$

Table 2

Analysis of Variance for Three Groups Differing in Degree of Antisocial Behavior

TAT Variables	Mean		Variance*	*F*	*p*
A. Fantasy Aggression					
Group I	6.08	Between Groups	136.69	23.98	.001
Group II	8.05	Within Groups	5.70		
Group III	11.64				
B. Number of 1 and 2 Remoteness Ratings / Fantasy Aggression					
Group I	.10	Between Groups	.71	35.50	.001
Group II	.17	Within Groups	.02		
Group III	.49				
C. External Punishment / Fantasy Aggression					
Group I	.87	Between Groups	.33	3.66	.05
Group II	.70	Within Groups	.09		
Group III	.62				
D. Internal Punishment / Fantasy Aggression					
Group I	.79	Between Groups	2.88	20.57	.001
Group II	.32	Within Groups	.14		
Group III	.08				
E. Internal Punishment / External Punishment					
Group I	1.03	Between Groups	3.63	12.96	.001
Group II	.52	Within Groups	.28		
Group III	.15				

* Between Groups $df = 2$. Within Groups $df = 54$.

.001. Since several authors have suggested the importance of social-class sanction as a factor governing the overt expression of covert aggressive needs, this variable was explored within the limits of the data. The bulk of the antisocial group was of lower-class background whereas the nonantisocial group was predominantly middle class (χ^2 significant at 5% level).

To press the examination further, the *Ss* in Groups I and III (least and most antisocial) were then separated according to social class. Using the *u* statistic described by Lord (3), the significance of the difference in number of aggressions for middle-class *Ss* only was calculated. The same was then done for lower-class members of Groups I and III. Group III *Ss* produced significantly more aggressive themes than Group I *Ss* regardless of social class. Although the number of *Ss* on whom these comparisons are based are relatively few (15 middle-class *Ss* in Group I and 5 in Group III, 6 lower-class *Ss* in Group I and 9 in Group III), the *p* values were .02 and .01. These several findings confirm the validity of Hypothesis 1 which concerns the relationship between aggression in fantasy and behavior.

To evaluate Hypothesis 2 the ratio of one and two remoteness ratings to total fantasy aggression score was compared for the three groups. The results, shown in Block B of Table 1, reveal the difference between the means to be significant well beyond $p = .001$. Not only does the antisocial individual respond with more aggressive themes, but the quality of his fantasy aggression is far more direct and undisguised than is true for the nonantisocial individual.

Block C of Table 1 compares the ratio of external punishment to fantasy aggression for the three groups. The significance of the difference between the means attains a value of $p = .05$. Thus an important force inhibiting the nonantisocial group appears to involve some sort of anticipated punitive action from without, giving support to Hypothesis 3.

Hypothesis 4 calls for a closer look at the fantasied external punishments. These were separated into three categories according to the following principles: (*a*) external punishment occurring subsequent to fantasy aggression and as a direct consequence of such aggression, e.g., "he was sent to jail for stealing"; (*b*) external punishment occurring prior to fantasy aggression and used as a justification for such aggression, e.g., "she was mean to him so he hit her"; (*c*) external punishment having neither of the above relationships to fantasy aggression, e.g., "she held him back because she was afraid he might get hurt." The total number of external punishments is larger for Group III than Group I with $p = .05$. It is clear, however, that TAT themes of external punishment are far more frequently utilized as excuses for aggression by antisocial members of Group III ($p = .001$), thus confirming Hypothesis 4. The mean number of themes classified in this manner was 3.64 for Group III, 1.25 for Group I. External punishment themes as a direct consequence of aggression are also more common among Group III *Ss* although the *t* value of 1.89 falls just short of significance. Since the number of postaggression punishment themes is obviously dependent upon the number of fantasy aggressions, an attempt was made to compare retaliatory punishment themes on a proportional basis. The data showed that the mean ratio of retaliatory punishment themes to fantasy aggression is identical for Groups I and III so that it was not even necessary to calculate a significance statistic. Interestingly enough, the number of external punishment themes unrelated to fantasy aggression is larger among the nonantisocial members of Group I ($p = .01$).

Block D of Table 1 indicates that the differences between the group means, with respect to the quantity of guilt and shame themes proportional to aggression, is significant at the value $p = .001$. Since the antisocial *Ss* produced significantly more aggressive themes than the nonantisocial *Ss*, one may wonder whether this alone is not the crucial determinant in creating differential values of the ratio of internal punishment to fantasy aggression. If one were able to find individuals in Groups I and III with aggressive impulses of similar strength, then the effect of internal punishment expectations could be independently evaluated. Overlapping in the number of fantasy aggressions between *Ss* of Groups I and III permitted comparison of six members from each group. They were the six highest and lowest in Groups I and III respectively in number of fantasy aggression themes. The *u* statistic shows no significant difference between the two groups with respect to fantasy aggression. However, the proportion of internal punishment to fantasy aggression is much larger for the nonantisocial group ($p = .001$). The mean ratio was .11 for the six Group III members and .54 for the six Group I members.

The relationship between socioeconomic class and antisocial rating raises a similar question regarding internal inhibition as was posed for fantasy aggression. Does the difference in

Table 3

t Ratios of Group Differences *

Groups compared	Fantasy aggression	Number of 1 & 2 Remoteness ratings / Fantasy aggression	External punishment / Fantasy aggression	Internal punishment / Fantasy aggression	Internal punishment / External punishment
Group I vs. Group II	3.18	1.75	1.77	3.36	2.83
Group II vs. Group III	4.17	5.33	1.09	6.00	4.07
Group I vs. Group III	6.04	7.84	2.27	4.44	4.40

* Minimum *t* ratios required for significance: 1.69 = <.10, 2.04 = <.05, 2.75 = <.01, 3.64 = <.001.

number of internal punishment themes for antisocial and nonantisocial groups hold true irrespective of social class membership? The *u* statistic shows clearly that Group I *Ss* produce significantly more internal punishment themes than Group III *Ss* regardless of social class. The difference between the middle-class members of Groups I and III is significant beyond the .001 level and among lower-class members of the two groups the difference achieves significance at the .01 level.

As shown in Block E of Table 1, the ratio of internal to external punishment themes differs significantly for the three groups ($p = .001$). The much higher ratio for Group I *Ss* suggests that internal punishment anticipations are a more potent deterrent to antisocial behavior than are external punishment anticipations, supporting Hypothesis 6.

Discriminative Power

Since a widely used diagnostic instrument was employed in this study it was deemed desirable to obtain some measure of diagnostic power or forecasting efficiency as distinguished from statistical significance. In order to evaluate discriminative power accurately it is necessary to make three comparisons of overlapping scores: Group I with Group II, Group II with Group III, and Group I with Group III. In each instance the median of the overlapping scores of the two populations being compared was selected as the cutoff point. Thus there were three different cutoff points for three separate comparisons. The percentage of correct classifications is calculated from the number of cases in each two-group comparison falling on the predicted side of the overlap median score. The data in Table 4 suggest that the most discriminating variables are the two internal punishment ratios. In clinical practice, however, it would seem appropriate to approach this classification problem by judging the interrelationship among the factors. For example, one of the members of Group I had a high fantasy aggression score but fell well within the nonantisocial group in all other scoring categories.

Table 4

Discriminative Power of Indicators Obtained from the TAT*

Variable	Groups I & II	Groups II & III	Groups I & III
Fantasy Aggression	68	67	76
Number of 1 & 2 Remoteness Scores / Fantasy Aggression	58	82	90
External Punishment / Fantasy Aggression	63	58	66
Internal Punishment / Fantasy Aggression	77	88	97
Internal Punishment / External Punishment	70	82	92

* Numbers in table refer to percentage of cases correctly classified.

Discussion

Since the finding of a positive relationship between fantasy and aggressive behavior held true for *Ss* of the middle as well as the lower class some attempt at explanation seems warranted in view of the contrary results of Sanford *et al.* (9) and Murray (5). The range in severity of aggressive behavior among *Ss* in the present study seems far greater than was true for either of the previous studies and this, of course, made for an easier discrimination. Furthermore, it is the author's impression, also suggested by Redl (8) and Feshbach (1), that certain individuals who characteristically behave in impulsive, antisocial fashion not only find fantasy an unsatisfactory substitute for direct action but may actually be stimulated and provoked to act out by a succession of onrushing hostile images.

The association between lower-class membership and antisocial behavior is perhaps partially reflective of the freer rein given aggressive impulses in this group as suggested by previous authors. It is important to note, however, that when antisocial behavior does occur, whether it be among middle- or lower-class *Ss*, these individuals manifest stronger aggressive impulses and/or a weaker internal control system than nonantisocial individuals. Thus the distinguishing feature is the indi-

vidual impulse-control balance, although specific cultural groups may tend to foster varying types of impulse-control balance with respect to aggression.

The analysis of remoteness of aggression is related to the problem of symbolic representation of impulses and conflicts in projective data. This investigation demonstrated that antisocial *Ss* were more direct in their expression of aggression than nonantisocial *Ss*. It appears, then, that defensive inhibitions alter not only the sheer quantity of aggressive themes but also the way in which they are formulated. This is a common enough clinical observation but one which does not seem to have been given sufficient weight in experiments concerned with the problem of inferring behavioral tendencies from projective material.

Anticipation of punishment as a consequence of aggressive behavior was no more prominent among nonantisocial *Ss* than antisocial *Ss*. This is not to say that such expectation may not serve to deter antisocial action in many instances. However, the results do suggest that it is not one of the critical distinctions between the two groups. On the other hand, a significantly larger number of punishment themes unrelated to aggression either as retaliation or justification was characteristic of Group I *Ss*. The content of themes classified in this manner suggests a greater awareness and acceptance of external direction and restraint. This finding may relate to the demonstration by Pastore (7) of the importance of perceived arbitrariness of the frustrating situation as a determinant of aggressive reaction.

As noted earlier in this paper there is a marked tendency for antisocial *Ss* to justify their aggressive reactions by pointing to the harsh treatment imposed on them by others. If antisocial individuals can prove that the world is against them, then there is no need to recognize any obligations to others. The very fact that they so frequently find it necessary to resort to this projection technique argues that the delinquent ego is busily engaged in disposing of still existing remnants of guilt. One of the implications of this state of affairs is that any general policy of retaliation only reinforces this defensive maneuver and makes it easier for the delinquent to pursue his antisocial course on a guilt-free basis. The basic importance of internal controls in the inhibition of antisocial behavior is amply demonstrated by the data. Certainly the evasion of guilt, shame, or remorse is a far more vital mark of the delinquent than any lack of awareness concerning potential external punishment consequent to antisocial behavior.

Another pertinent topic involves the hypothesis that fantasy behavior has a compensatory function. Tomkins (10) assumes this hypothesis in stating that TAT aggressive fantasies are found more commonly among neurotics and normals than antisocial characters. The results of the present study suggest that the opposite is the empirical fact, given *Ss* differing widely in antisocial behavior. However, this finding is not necessarily contradictory to the fantasy substitution hypothesis. One possibility is that the more potent and numerous aggressive impulses of the antisocial *Ss* obscure any relative increment in fantasy aggression in the nonantisocial *Ss*. An alternative suggestion mentioned earlier is that fantasy is not universally employed as a defensive or adaptive measure. Clinical experience tells us that individuals differ widely in the use of fantasy and it may simply be that many directly hostile individuals have not learned to resort to fantasy for substitute gratification.

Summary

This study has relevance to two problems. Its first interest was to examine certain personality variables, particularly in terms of the impulse-control balance, in order to establish their relationship to overt aggressive tendencies. The TAT was the instrument through which information was gathered about the nature of and interrelationships between aggressive impulses, anticipated internal punishment, and anticipated external punishment. The second major issue concerns the relationship between fantasy and reality behavior and the task of formulating from projective data alone the various cues which might permit differential predictions about behavior tendencies insofar as overt expression is concerned.

Fifty-seven young adult soldiers, the large majority of them basic trainees and all of them psychiatric referrals to the Mental Hygiene Clinic, served as *Ss*. On the basis of social history data covering aggressive and antisocial behavioral manifestations in both civilian and military life, each *S* was rated as to overt antisocial tendencies. It was found that antisocial *Ss* differ from nonantisocial *Ss* in quantity and quality of fantasy aggressions, the former producing not only a larger number of fantasy aggressions but also finding it less necessary to obscure or minimize the hostile impulse. The superego variable as measured by anticipated internal punishments was found to be of far greater significance in inhibiting antisocial behavior than fear of retaliatory punishment which did not even differentiate between groups. An interesting finding dealt with the tendency of antisocial *Ss* to utilize themes of external punishment as a justification for their fantasy aggressions. Some of the implications of these results were discussed. An effort was made to deal with the problem of socioeconomic class factors as they might relate to the data. Finally, emphasis was placed on the fact that, for behavioral prediction from projective material, the same attention must be given to repressive, inhibitory forces manifested in the data as is given to the impulse or drive system.

Received January 25, 1956.

References

1. Feshbach, S. The drive-reducing functions of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, **50**, 3–11.
2. Lindzey, G. Thematic Apperception Test; interpretive assumptions and related empirical evidence. *Psychol. Bull.*, 1952, **49**, 1–25.
3. Lord, E. The use of range in place of standard deviation in the *t*-test. *Biometrika*, 1947, **34**, 41–67.
4. McNemar, Q. *Psychological statistics*. New York: Wiley, 1949.
5. Murray, H. A. *Thematic Apperception Test manual*. Cambridge: Harvard Univer. Press, 1943.
6. Mussen, P. H., & Naylor, H. K. Relationships between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, **49**, 235–240.
7. Pastore, M. The role of arbitrariness in the frustration-aggression hypothesis. *J. abnorm. soc. Psychol.*, 1952, **47**, 728–731.
8. Redl, F., & Wineman, D. *Children who hate*. Glencoe, Ill.: The Free Press, 1951.
9. Sanford, R. N., Adkins, M. M., Miller, R. B., Cobb, E. A., and others. Physique, personality, and scholarship: A cooperative study of school children. *Monogr. Soc. Res. Child Develpm.*, 1943, **8**, No. 1.
10. Tomkins, S. S. *The Thematic Apperception Test*. New York: Grune & Stratton, 1947.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Situational Influences on Rorschach Responses

Edith M. Henry and Julian B. Rotter
The Ohio State University

The widespread use of "projective" tests for purposes of diagnosis and prediction of behavior has stimulated extensive research investigating the validity of the underlying assumptions of these tests. One of the most significant areas in which research has raised questions is that of the effect of immediate influences on projective test productions. In social learning theory (14, 15) such influences are conceptualized as the effect of the psychological situation on behavior.

A growing body of literature demonstrates that many factors other than an individual's basic personality structure are operating in any test situation, and that these factors are important determiners of the types of responses to be given. That is, individuals adapt their behavior to the total environment in which they find themselves, and external stimuli as well as inner motivations and dynamics shape behavior at any given time. The present study is an attempt to show further the importance of situational factors and to predict the types of changes in response which will occur as a result of a particular type of change in the testing situation utilizing the Rorschach Test.

More specifically this study is concerned with the subject's (*S*'s) belief regarding the purpose of the test as it affects his test responses. A number of previous investigations (3, 4, 7, 9, 10, 11, 12, 16) have dealt with other kinds of situational influences on the Rorschach test, including such variables as the effect of immediate stress, the sex and personality of the examiner, suggestion, and the physical place of testing. There have also been a number of studies more closely related to the present one.

Fosberg (5, 6) found little consistent effect of asking *Ss* to make the best possible and the worst impression without, however, clarifying how they should do this. Carp and Shavsin (2) however in a similar study found intra-individual differences although there were no statistical group differences. That is, *Ss* changed their responses considerably but they did not do so in a consistent manner. Luchins (13) in investigating why *Ss* did not give certain kinds of responses in the early part of the Rorschach test found that one of the more frequent reasons was the *S*'s fear of being labeled "psycho." Other reasons also implied that the *S* was responding to his interpretation of the purpose of the test. Similarly, Hutt, Gibby, Milton, and Pottharst (8) have shown that variations in instructions will produce significant variations in responses. Schachtel (17) has discussed in detail how the *S*'s implicit definition of the Rorschach test will affect responses but has not investigated these experimentally.

In the present study, the situational variable of the *S*'s definition of the test was manipulated and then investigated. It was felt that one of the most widespread views of the Rorschach, resulting from publicity of mass media, is that the Rorschach test is a measure of "sanity," that this test will uncover imbalances in a person's emotional life. In accord with the previous discussion, it was hypothesized that such a view of the test would alter an individual's responses in the direction of caution and the giving of only what most *Ss* would consider to be "safe" responses. In order to test these assumptions, the experimental *Ss* were given instructions about the test which included the suggestion that this test is designed to discover emotional disturbances of a serious nature in mental patients. An underlying assumption of the study was that, when *Ss* respond in a particular way after being given a hypothesis about the test, these responses are comparable to those of another individual who holds the same hypothesis implicitly. The following hypotheses were formulated:

The experimental group will give significantly; (*a*) fewer responses in total (*R*) than will the control group; (*b*) more responses scored as form (*F*) than will the control group; (*c*) more responses of good form-level (*F* +) than will the control groups; (*d*) more large detail (*D*) responses than will the control group; (*e*) fewer whole (*W*) responses than the control group; (*f*) more popular (*P*) responses than will the control group; (*g*) more animal (*A*) responses than will the control group; (*h*) fewer responses of an aggressive and violent nature than will the control group.

In all these hypotheses the assumption is made that *S*s in the experimental situation will give, in the main, those responses which they consider to be "safe," inhibiting those which they consider suspect, those which would tend to place them in an unfavorable light. That is, they will give more responses which are conforming, which they expect to be similar to what most other people would see on the same blot or area. The present study differs from other Rorschach studies in that a situational variable is manipulated, and hypotheses are formulated about the changes in the determinants on a logical, "common-sense" basis rather than on the basis of the usual theoretical meaning of the various determinants although the two approaches coincided in some instances. The hypothesis about decrease in total number of responses for the experimental *S*s, for instance, was formulated because of the general tendency toward caution and inhibition. Similarly, an increase in pure form responses was hypothesized not because of its theoretical meaning as a measure of intellectual control but rather because form responses are often easier to see and those based on color, texture, or movement are frequently less clear and would be considered suspect. Large detail responses, again, were believed to correspond to the more obvious percepts; the preponderance of popular responses is in line with a "common-sense" attempt at caution and safety, at seeing things as most other people are likely to see them. The hypothesized increase in animal responses concurs with an increase in popular responses. The assumption that a large number of animal responses suggests intellectual immaturity was not employed. It was felt that a reduction in aggressive and violent types of responses would result from attempts by *S*s to make their responses as innocuous as possible, to inhibit any signs of lack of control over feelings.

Procedure

Subjects. Sixty female college students in elementary psychology were randomly allotted to two groups of thirty.

Procedure. When *S* entered the experimental room, she was asked the following initial questions:

> "Have you ever heard of the Rorschach test . . . the inkblot test?" If the response was affirmative, the following question was asked:
>
> "Tell me what you remember about it? . . . What do you think it does? What is it for?"

Responses to these questions were recorded by *E* who then gave the main instructions. For the experimental group these were:

> "This is a series of cards with inkblots on them. People look at the cards and tell what the blots remind them of. For many years the test has been used in mental hospitals to study emotional disturbances and to discover how seriously disturbed patients were. Now we are making a college survey with it.
>
> "Now I want you to look at the blots and tell me what they could be, what they look like to you, or what they remind you of. People see all different things in the blots—there are no right or wrong answers."

The individuals in the control group, every second *S*, were also asked the questions about their previous acquaintance with the test, and then were given the standard Klopfer instructions:

> "This is a series of cards with inkblots on them. I want you to look at them and tell me what they could be, what they look like to you, or what they remind you of. People see all different things in the blots—there are no right or wrong answers."

Following this, the Rorschach was administered in the standard manner to all *S*s, and a fairly thorough inquiry followed.

After the test, the experimental *S*s were asked a number of general questions, designed to elicit information about the effectiveness of the instructions.

Scoring. All identification was removed from the records, and the protocols were scored according to Beck's (1) manual. An independent rater then scored twelve records, also using Beck's criteria. A short scoring manual was constructed for the presence of content involving aggression or violence. The categories used were as follows: explosions; fire; blood; people or animals battered, beaten, or demolished; fighting or killing; monsters involved in some kind of threatening action. A frequency count of these themes was then made by the experimenter and by the same independent rater.[1]

Statistical methods. To test reliability, a percentage of agreement was computed between the two scorers on twelve records. A Pearson product-moment correlation was used to test reliability of scoring for aggressive content on all sixty protocols.

To test the hypothesis dealing with total number of responses, a *t* test for differences between means was used. For all other hypotheses, two different techniques were used. The first technique involved converting all scores to percentages and then computing *t* tests for differences between the means of the percentages. The second technique was essentially a method of equalizing the scores to account for any differences in productivity (Total *R*) between the two groups.[2]

[1] The authors would like gratefully to acknowledge the assistance of Mr. Walter Mischel in this respect.

[2] We are grateful to Dr. Robert Wherry who suggested a method for equalizing the effects of an unequal number of responses in the two groups. To equalize scores the following steps were taken for each variable except *R:* (*a*) The total number of responses for both experimental and control groups was computed; (*b*) The total number of responses on any particular variable, such as *F*, were computed, and this number was divided by the total *R* to give a predicted score on the variable under consideration for the group as a whole; (*c*) To arrive at an individual predicted score, this total predicted score was multiplied by each individual's total *R;* (*d*) This predicted score was then subtracted from each individual's actual score on the variable being considered. This score served as an indication of where each individual's score fell in relation to the score that would be predicted from total group performance. These scores were then used in the computation of *t* tests of significance between the means of the two groups.

Results

Reliability of measures. Percentages of agreement were calculated between the scoring of *E* and an independent rate on twelve protocols. (On the remaining protocols, the *E*'s scoring alone was used with all identification first removed from the records.) Separate computations were performed for location, determinants, and content. The percentage of agreement for scoring of location was .95, for determinants, .82, and for content, .92. A second reliability check concerned the scoring for aggressive and violent content, scoring having been based on a brief manual constructed by *E*. Here, all sixty protocols were used. Pearson product-moment correlations were used here, with separate computations for the experimental and control groups. The correlation coefficient for the experimental group was .83, and for the control group, .66.

The results of the tests of the experimental hypotheses are given in Table 1.

a. R (total responses) lower in experimental than control group: The experimental group mean was 16.07, and the control group mean was 23.47. A *t* of 4.11 supports the hypothesis beyond the .001 level of confidence.

Table 1

Sigmas, Means, *t*s and Significance of Differences Between Means for the 30 Experimental and 30 Control Subjects

Variable	σ_e	Mean$_e$	σ_c	Mean$_c$	*t*	*p*
Responses	6.20	16.07	7.52	23.47	4.11	<.001
From Equalized Scores						
Popular	1.85	.56	1.97	−.56	2.23	<.05
Form	2.43	.76	3.56	−.75	1.88	<.10
Form-Level	1.17	.66	2.49	−.67	2.62	<.02
Aggressive	1.94	−.48	1.80	−.12	1.28	<.20
Animal	2.14	.73	2.86	−.74	2.25	<.05
Whole	3.48	.15	5.04	.87	.63	<.60
Detail	3.91	.23	3.97	.22	.02	<.90
From Percentages						
Popular	.12	.32	.04	.26	2.29	<.02
Form	.19	.72	.13	.61	2.50	<.02
Form-Level	.08	.87	.08	.80	2.70	<.01
Aggressive	.11	.07	.10	.10	1.08	<.20
Animal	.14	.53	.12	.43	2.93	<.01
Whole	.23	.45	.46	.44	.92	<.40
Detail	.24	.50	.70	.49	.40	<.70

b. More *F* (form) responses in experimental than control group: Using equalized *R* scores, means of .76 and − .75 were obtained for the experimental and control groups, respectively. A *t* of 1.88 was obtained, which yields a *p* value between the .05 and .10 levels and tends to substantiate the hypothesis. Using percentage scores, the mean value for the experimental group was .72 and .61 for the control group. A *t* of 2.5 supports the hypothesis between the .02 and .01 levels of confidence.

c. More *F* + (good form level) responses in experimental than control group: With equalized scores, means were .66 and − .67 for experimental and control groups, respectively. A *t* of 2.62 supports the hypothesis between the .02 and .01 levels. A *t* calculated from percentages yielded a value of 2.70, significant beyond the .01 level in the predicted direction. Mean values were .86 and .80 for the experimental group and the control group, respectively.

d. More *D* (large detail) responses in experimental than control group: With equalized scores, means of .23 and .22 were obtained for the experimental and control groups. A *t* of .02 and, with percentages, a *t* of .34 failed to give any substantiation to this hypothesis. Means of percentage scores were, respectively, .50 and .49.

e. Fewer *W* (whole) responses in experimental than control group: Using equalized scores, a mean of .15 was obtained for the experimental group and of .87 for the control group. A *t* of .63 with equalized scores and .92 with percentage scores did not substantiate this hypothesis. Means with percentage scores were .45 and .44 respectively.

f. More *P* (popular) responses in experimental than control group: Experimental group mean, with equalized scores, was .56; control group mean was − .56. A *t* of 2.23 with equalized scores lent support to this hypothesis between the .02 and .05 level. With percentages, means were .32 and .26, respectively. A *t* of 2.29 supported the hypothesis between the .025 and .01 levels.

g. More *A* (animal) responses in experimental than control group: Using equalized *R* scores, means were .72 and − .74, respectively. A *t* of 2.25 supports this hypothesis between the .02 and .01 levels. With percentages, means were .53 and .43 for experimental and control groups, respectively. With *t* calculated from percentage scores, significance was between the .01 and .001 levels.

h. Fewer aggressive responses in experimental than control group: With equalized scores, means were − .48 and − .12, respectively. A *t* of 1.28 fell between the .10 and .20 levels of confidence. Using percentage scores, means were .07 and .10, respectively. A *t* of 1.08 also fell between the .10 and .20 levels, and, therefore, did not significantly substantiate this hypothesis but was in the right direction.

Discussion

The results of this study, in general, seem to substantiate the hypotheses advanced. In regard to aggressive responses, the total number produced in the group was too few to obtain significance in the results and the assumption that the more conforming responses were frequently more *D* than *W* was not borne out. The other hypotheses regarding the number of responses, the number of form responses, form-level, popular responses, and animal responses, were all supported to a highly significant degree.

The real significance of this study does not lie in its demonstration of the effect of a different type of instructions but rather in the implications that subjective hypotheses about what the test is testing will have a significant effect on the *S*'s responses. As applied to the Rorschach test some specific information is presented in this study as to what that effect may be when the *S* has the hypothesis or belief that the test is essentially a test of insanity, mental disturbance, psychopathology, etc. Such a belief is probably fairly commonly held as a result of the widespread popularization of the Rorschach through magazine articles, Sunday supplements, movies, radio, television, and also through direct communication. The results support strongly Schachtel's assertion that it is extremely important for an examiner to find out what these implicit beliefs regarding the test are, and to evaluate test results in the light of them.

It is of some interest that the general questioning of the *Ss* following the test administration did not reveal too much in the way of awareness that the *Ss* were trying to appear in a good light, not mentally disturbed, etc. Some *Ss* after considerable pushing would finally agree that they tried to avoid mentioning anything that might sound peculiar, but in general, more prediction could probably be obtained by assuming certain effects likely to occur as a result of the *S*'s hypotheses about the test purposes than could be obtained by questioning the *S* regarding his motives in taking the test.

The study also tends to support, at least by implication, the general notion that conceptualizing the *S*'s behavior (including test responses) in a test situation as an attempt to reach some goal to give a particular kind of impression or to defend himself against some threat can be a useful predictive method from a common-sense point of view.

The results also tend to support the now growing body of literature which emphasizes the need to interpret psychological test results in light of a host of situational variables, no matter how indirect the test, rather than limiting analysis to the test protocols or profiles obtained from several tests or subtests.

Summary

This study measured the influence of one aspect of the psychological situation, knowledge of test purpose, on responses to the Rorschach test. The *Ss* were sixty female students in elementary psychology. Experimental *Ss* were told that the Rorschach is used to discover serious emotional disturbances, then given standard Klopfer instructions; control *Ss* received only Klopfer instructions. Hypotheses were that experimental *Ss* would give fewer total, aggressive, and whole responses; more pure form, good form-level, popular, animal, and large detail responses. The specific hypotheses all involved the general hypothesis that experimental *Ss* would be cautious and conforming. Results were analyzed by two methods for every variable except total responses. By one, raw scores were converted into percentages, and *t* tests computed. The second involved a technique of equalizing scores to eliminate any influence of reduced number of total responses in one group. With total responses, *t*'s were computed on differences between means. Results showed fewer responses in experimental group (beyond the 1% level), more good form-level, popular, and animal responses (between 5% and 1% levels), more form responses (between 10% and 5% levels), fewer aggressive responses (between 20% and 10% levels). Location did not vary significantly between groups.

The results tend to support the general hypotheses that: (*a*) a knowledge of the *S*'s beliefs regarding the purposes of the test should be an important aspect of the interpretation of the test, (*b*) a "common-sense" analysis of the *S*'s approach to the whole testing situation would be a fruitful part of the test interpretation, and (*c*) situational determinants significantly influence test responses.

Received January 16, 1956.

References

1. Beck, S. *Rorschach's test.* Vol. 1. *Elementary principles.* New York: Grune & Stratton, 1944.
2. Carp, A. L., & Shavzin, A. R. The susceptibility to falsification of the Rorschach psychodiagnostic technique. *J. consult. Psychol.*, 1950, **14**, 230–233.
3. Coffin, T. E. Some conditions of suggestion and suggestibility. *Psychol. Monogr.*, 1941, **53**, No. 4 (Whole No. 241).
4. Eichler, R. M. Experimental stress and alleged Rorschach indices of anxiety. *J. abnorm. soc. Psychol.*, 1951, **46**, 344–355.
5. Fosberg, I. A. Rorschach reactions under varied instructions. *Rorschach Res. Exch.*, 1938, **3**, 12–31.
6. Fosberg, I. A. How do subjects fake results on the Rorschach test? *Rorschach Res. Exch.*, 1943, **7**, 119–121.
7. Gibby, R. C., Miller, D. R., & Walker, E. L. The examiner's influence on the Rorschach protocol. *J. consult. Psychol.*, 1953, **17**, 425–428.
8. Hutt, M. L., Gibby, R., Milton, E. O., & Pottharst, K. The effect of varied experimental sets upon Rorschach test performance. *J. proj. Tech.*, 1950, **14**, 181–187.
9. Kimble, G. A. Social influence on Rorschach records. *J. abnorm. soc. Psychol.*, 1945, **40**, 89–93.
10. Klatskin, E. H. An analysis of the effect of the

test situation upon the Rorschach records. *J. proj. Tech.*, 1952, **16**, 193–199.

11. Levine, K. N., Grassi, J. R., & Gerson, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach, a case study. *Rorschach Res. Exch.*, 1943, **7**, 130–144.
12. Lord, Edith. Experimentally induced variations in Rorschach performance. *Psychol. Monogr.*, 1950, **64**, No. 10 (Whole No. 316).
13. Luchins, A. J. Situational and attitudinal influences on Rorschach responses. *Amer. J. Psychiat.*, 1947, **103**, 780–784.
14. Rotter, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.
15. Rotter, J. B. The role of the psychological situation in determining the direction of human behavior. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln: Univer. of Nebraska Press, 1955. Pp. 245–268.
16. Sarbin, T. R. Rorschach patterns under hypnosis. *Amer. J. Orthopsychiat.*, 1939, **9**, 315–318.
17. Schachtel, E. G. Subjective definitions of the Rorschach Test situation and their effect on test performance. *Psychiatry*, 1945, **8**, 419–448.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Affective Control and Rorschach Human Movement Responses[1]

Julian Meltzoff and Dorothy Litwin[2]
Veterans Administration Regional Office, Brooklyn, New York

Previous studies have demonstrated that one of the important behavioral correlates of Rorschach Human Movement Responses (*M*) is the ability to inhibit or delay motor acts voluntarily. In support of hypotheses drawn from both Werner's Sensory-Tonic theory and Rorschach's empirical observations, a series of experiments (3, 4, 5, 6) has shown that *Ss* who are responsive to the kinesthetic aspects of the inkblots are better able to inhibit motor acts than those who are not. Further investigations (1, 2) have extended this relationship to the cognitive realm. Experiments by Meltzoff and Levine have shown that good motor inhibition goes hand in hand with the ability to inhibit learned associations voluntarily and substitute new ones. Furthermore, effective cognitive inhibition was shown to be positively related to the perception of movement in the same way as was motor inhibition. Thus *M* appears to be an indicator of the ability to inhibit or control both motor acts and thoughts.

The aim of the present study is to extend the relationship of *M* and inhibitory ability one step further to include the control of feelings as well.

The Experiment

Hypothesis. A direct relationship between the production of Rorschach Human Movement Responses and affective control is hypothesized. It is predicted that more *Ss* with high *M* than with low *M* will be able to inhibit the overt expression of an experimentally induced affect.

Procedure. The general method consists of inducing an affect in the *Ss*, instructing them *not* to express their feelings, and assessing the success or failure of their control. Consideration of the various methodological and ethical problems involved in experimenting with strong, unpleasant affects, led to the selection of one of the pleasant emotions instead. An effort was made to induce an affective state by presenting a humorous stimulus. The stimulus chosen was "The Jones Laughing Record," a recording by Spike Jones and his orchestra for RCA-Victor. This is a recording of contagious laughter with an accompaniment of humorous background music and assorted noises.

The judgment as to what constituted an affective response was made on the basis of fairly gross pre-established criteria. Any reaction ranging from the first sign of a smile to an outburst of laughter was considered as an affective response. This included any reaction from a break in the set of the mouth with a widening and upturning at the corners, to the opening of the mouth accompanied by the characteristic spasmodic inspiration and expiration of laughter.

In a preliminary trial with two observers there was a complete agreement on judgments of the mere presence or absence of an affective response at some point during the 2½-minute recording. A single observer was used in the experiment proper.

The recording was pretested on a sample of 20 mixed, haphazardly selected college students, 19 of whom responded by smiling or laughing. Following the record, *Ss* were asked

[1] Paper read at East. Psychol. Ass., Atlantic City, March, 1956.

[2] From the Veterans Administration Regional Office, Brooklyn, New York.

to respond in the usual manner to Cards III and VII of the Rorschach. The sample was then divided at the median into High (2 or more) and Low (less than 2) *M* groups. It was determined that there was no significant difference between the two groups in the number of *Ss* who responded with an overt expression of affect ($\chi^2 = .12$, $p = .70$). Thus the record was successful in inducing an impulse to laugh, and the expression of that impulse, regardless of the amount of *M* produced by the *Ss*.

Sixty-eight additional *Ss* from the same general population were introduced to the same experimental procedure with one crucial modification. Prior to hearing the record, *Ss* were instructed *not* to smile or laugh. This sample was similarly divided at the median into high *M* and low *M* groups. The two groups were compared on their success or failure in controlling the expression of the experimentally induced affect.

It is to be noted that the Rorschach was administered *following* the laugh record in order to rule out any bias in the judgment of affective response. Thus the experimenter had no way of knowing whether a subject was going to fall in the high *M* or low *M* group at the time of making the judgment about the presence or absence of an affective response.

Rorschach protocols were scored independently by two judges without knowledge of the *Ss* performance on the inhibition task. In the few cases of scoring disagreement the judgment of a third psychologist served as the deciding factor.

Results

Results were in the predicted direction. Of the 34 *Ss* in the high *M* group, 20 were able to listen to the entire record without any observable display of feeling. Only 11 of the 34 *Ss* in the low *M* group were similarly successful (Table 1).

Table 1

Affective Control of High and Low *M* Groups

Group	Successful Control	Failure to Control
High *M* (2 or more)	20	14
Low *M* (less than 2)	11	23

Since the direction was hypothesized, a one-tailed test of significance was used. The obtained chi square of 3.79 is significant at the .03 level. Analysis of the data in terms of the total number of responses (*R*) reveals that the obtained relationship is not dependent upon *R* ($\chi^2 = 1.27$, $p > .20$). Tetrachoric correlation between *M* and affective control was .42 ($p < .01$).

Discussion

These results support the hypothesis of a significant positive relationship between *M* and the voluntary inhibition of expression of an experimentally induced impulse to laugh. Although these data cannot be generalized to the control of other affects, one could hypothesize that further experimentation with other affects would reveal similar relationships.

It is of interest to speculate about the processes involved in controlling affective expression. Is there any fundamental difference in process between the inhibition of the *expression* of an affect and the control of the affect itself? The two may be equivalent only from the point of view of the outside observer. Controlling affective expression appears to have both motor and cognitive components. A person can keep from expressing an impulse to laugh, for example, by controlling certain facial muscles. The affect can remain even though its overt expression is controlled by motor means. Some of the subjects reported that they felt very much like laughing, but restrained themselves by clenching their teeth and compressing their lips.

The other way of controlling the expression of affect is through the modification of the affective tone itself by cognitive processes which alter the autonomic reaction pattern. By way of illustration, one of the subjects reported that she kept from laughing by thinking of her mother dying. She was thereby able to change the affective tone through cognitive means. The affect need not be changed to the extent of substituting an opposite one, as thoughts could be used in a variety of ways merely to neutralize the existing feel-

ing. We have already cited experiments which have demonstrated that the abilities to inhibit motorically and cognitively are each correlated with *M*. If voluntary control of the impulse to express an affect is viewed in terms of its motoric and cognitive components, the present experiment may simply be reflecting these relationships in a more complex setting.

As suggested above, there may be two basic styles of inhibiting affective expression—one primarily motoric, and the other principally cognitive. We might speculate further and invoke a concept of motor-cognitive vicariousness. This would point up the possibility that motor inhibition tends to lead to the production of thoughts or fantasies which in turn serve to modify the affective tone. Observations and reports from the *Ss* suggest that their initial adaptation to the task was through motor control. It appeared that this was later supplemented by cognitive methods at different stages according to the individual. This progression from motor to cognitive control can be thought of as a recapitulation of the developmental process, in that the ability to use cognitive processes to control feelings might be expected to appear at a later stage of personality development. Observations of young children suggest that control largely through motor means is gradually supplemented by the acquisition of a variety of cognitive techniques which can modify the affective state. When these mechanisms are refined in the normal adult, it is more difficult for the observer to detect the affective state than it is in the emotionally labile child. We would expect different behavorial and symptomatic manifestations in adults with these two different styles of control in accordance with the differences in their chronic reaction pattern. Thus we might anticipate different symptom complexes in individuals who are chronically exerting motor control in contrast to others whose control is exercised largely through thinking. It would therefore be important to contrast individuals who are found to utilize predominantly one or the other of these methods from the point of view of relating their style of control to developmental levels and to the complicated problem of symptom choice in psychopathology.

Summary

Previous experiments have demonstrated motor and cognitive inhibition to be behavioral correlates of Rorschach Human Movement responses. The present study extends the relationship of *M* and inhibition ability further to include the control of affective impulses as well. An impulse to laugh was induced, and *Ss* instructed to refrain from any affective expression. In support of the hypothesis, more high *M* than low *M* subjects were successful at inhibiting any overt expression of the experimentally induced affect.

Received February 21, 1956.

References

1. Levine, M., & Meltzoff, J. Cognitive inhibition and Rorschach human movement responses. *J. consult. Psychol.*, 1956, **20**, 119–122.
2. Meltzoff, J., & Levine, M. The relationship between motor and cognitive inhibition. *J. consult. Psychol.*, 1954, **18**, 355–358.
3. Meltzoff, J., Singer, J. L., & Korchin, S. J. Motor inhibition and Rorschach movement responses. *J. Pers.*, 1953, **21**, 400–410.
4. Singer, J. L., & Herman, J. Motor and fantasy correlates of Rorschach human movement responses. *J. consult. Psychol.*, 1954, **18**, 325–331.
5. Singer, J. L., Meltzoff, J., & Goldman, G. D. Rorschach movement responses following motor inhibition and hyperactivity. *J. consult. Psychol.*, 1952, **16**, 359–364.
6. Singer, J. L., & Spohn, H. E. Some behavioral correlates of Rorschach's experience type. *J. consult. Psychol.*, 1954, **18**, 1–9.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Use of the Kaleidoscope as an Adjunct to the Rorschach

Emanuel Starer
VA Hospital, Northport, N. Y.

One of the problems encountered in the psychological testing of many chronic psychotic patients is that they frequently produce only ten or fewer responses on the Rorschach. As a result, the examiner is hampered in his attempt to describe the underlying personality structure in detail.

The writer has employed the kaleidoscope with 20 chronic patients, with the instructions to report whatever was seen. After the patient had finished associating to a given stimulus, he was instructed to shift the kaleidoscope which would automatically introduce another stimulus. In this manner an unusually large number of stimuli could be introduced in a relatively short space of time. Approximately 15 to 20 minutes were allowed for each patient. The associations to the kaleidoscope could be analyzed in a manner similar to content analysis on the Rorschach.

Most of the patients produced two to three times the number of responses obtained on the Rorschach. Indeed, one patient who refused the Rorschach volunteered to associate to the kaleidoscopic stimuli and produced 30 responses which were helpful in evaluating some of his areas of conflict.

Examples of some of the responses obtained with the kaleidoscope include those frequently obtained from the Rorschach such as: bugs, butterflies, bats, insects, plants, flowers, and designs, usually geometric in nature.

Responses containing a color component included: a bloody rag, two big red blotches, face of a cat with red eyes and red nose, and I see a discolored ball, sort of pale and brown.

Some of the responses containing human, animal, or inanimate movement included: man's face, sort of stupid looking with sharp nose, statue of the Lord, He's bowing His head, face of a cat who is ready to eat you up, crabs clawing an insect, a spider making up his mind whether to poke at me or not, three bolts of lightning hitting an object from three different directions.

Finally, some examples of highly personalized associations included: pair of ice tongs like the ones I used when a boy, type of Pennsylvania-Dutch design, a hex sign, insignia of the devil, fundus of the uterus, mop of hair of a child, pancake hat for young lady, and motor neurone.

The kaleidoscope technique is recommended only when there is a paucity of associations on the Rorschach, and so can be considered only an adjunct. Much valuable information is lost since only a content analysis can be made. Perhaps the most objectionable to many investigators is that the stimuli are not constant and so can not be reduplicated. In spite of these and other objections, this examiner has found the adjunctive use of the kaleidoscope in personality evaluation to be worthy of further investigation and use in practical clinical testing situations, particularly with chronic psychotic patients.

Received February 29, 1956.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Use of Human Figure Drawings in the Diagnosis of Organic Pathology

Marvin Reznikoff and Donald Tomblen
The Institute of Living

Clinical psychologists are often asked to aid in the differential diagnosis of organic pathology. One of the widely used techniques in psychodiagnostics is the human figure-drawing test. Experience has suggested that this relatively unstandardized method may prove to be useful in attempting to evaluate the presence or absence of organic difficulties, but there are relatively few published studies in this area. Michal-Smith (2) reported the results of the H-T-P technique, comparing boys whose EEGs were abnormal with boys who were found to be normal with reference to their EEGs. Only one sign yielded a significant difference, namely, "line quality."

Machover (1) and Vernier (3) both listed indicators which they mentioned as "often seen" in the organic case. For Machover, these included: disproportionately large head; empty, large, poorly proportioned figures; weakly synthesized drawings; and simple, heavy line quality. Vernier suggested "petal" or scribbled fingers; center-of-page placement; shrunken arms and legs; erasures in arm and hand area; and eyes drawn as slits. In agreement with Machover, she also listed largeness of head and whole figure; emptiness, or lack of detail; weak synthesis; and poor proportions. However, she described the line quality as sketchy and/or broken, rather than heavy and simple.

The present study was undertaken to investigate: (*a*) the degree to which these various suggested indicators are actually more characteristic of the drawings of organic patients than they are of the drawings of other psychiatric groups, and (*b*) the extent to which these indicators may reliably be employed in the differential diagnosis of organicity.

Procedure

Sample. Three groups of 25 patients each were selected from the files on the basis of final diagnoses, which in this institution follow the criteria and classifications described in the *Diagnostic and Statistical Manual of Mental Disorders* prepared by the American Psychiatric Association. In one group, all patients had been given organic diagnoses, specifically "chronic brain syndrome." Another group of patients were diagnosed as schizophrenic, and the last group as having neurotic disorders. Organic findings were entirely absent in the schizophrenic and neurotic groups. All patients were at least average in intelligence and between 19 and 55 years of age, with a mean age of 35.4 years. All patients had drawn human figures during routine psychological testing. The patients had not been tested by either author, nor was any information about the personal history, diagnosis, or results of other psychological tests available to the authors when they scored the drawings.

Scoring method. A list of 17 indicators was compiled, based upon those suggested in the literature plus several others derived from the personal experience of the senior author. Scoring criteria for each indicator were delineated. In the instance of No. 15, "Enlarged figure size," actual measurement of the drawing length was recorded. For all other indicators, a simple plus-minus system was used. If an indicator was clearly present, it was scored plus; otherwise, it received a minus score. The indicators were:

Table 1

Analysis of Variance of Drawing Size Measurements of Organic, Schizophrenic, and Neurotic Patients

Source of variation	*df*	Mean square	*F*
Between groups	2	6.24	1.15
Within groups	72	5.44	
Total	74		

1. Emptiness of figure (lack of details)
2. Weak synthesis (parts loosely joined)
3. Parts misplaced (noticeably shifted)
4. Disproportionately large head
5. Shrunken arms and legs (thinner and shorter)
6. Other parts distorted (inappropriate size and/or shape)
7. "Petal" or scribbled fingers
8. Eyes drawn as slits
9. Center-of-page placement
10. Erasures in arm and hand area
11. Erasures elsewhere (two or more in single drawing)
12. Heavy, simple line quality
13. Sketchy line quality
14. Broken line quality
15. Enlarged figure size
16. Unlifelike pose
17. Erasures with only minimal improvement

The drawings of all 75 subjects were randomly arranged with all identifying data removed. Each of the authors independently scored every drawing for every indicator.

Table 2

Drawing Indicators Showing Significant Differences Between Groups of Organic and Neurotic Patients of Twenty-five Each

Indicator	Frequency of occurrence: Organic	Frequency of occurrence: Neurotic	Chi square	*p*
Weak synthesis	15	1	15.53	.001
Parts misplaced	15	1	15.53	.001
Shrunken arms and legs	14	2	11.12	.001
Parts other than head and extremities distorted	15	3	10.50	.01
Petal or scribbled fingers	13	4	5.70	.02
Emptiness of figures	15	5	6.75	.01

Two indicators proved to be unscorable in practice and therefore were eliminated. These were: No. 16, "Unlifelike pose," and No. 17, "Erasures with only minimal improvement," both original with the authors. It was impossible to arrive at a satisfactory, objective definition of unlifelike pose, and it was likewise extremely difficult to determine the degree of improvement after erasures without actually observing the drawing process itself. On the 15 retained indicators scored for every drawing, there was 91 per cent interscorer agreement. Minor differences of scoring were readily resolved in conference, still independent of any knowledge of the subject's diagnostic category.

Treatment of data. Indicator No. 15, "Enlarged figure size," was tested separately by means of an analysis of variance of figure measurements. All other indicators (Nos. 1–14) were studied together. A tabulation was made of the frequency of occurrence of each indicator among the 25 patients in each diagnostic group. Maximum frequency would thus be 25. Each group was then compared with the other two groups by means of chi square.

Analysis of variance was again utilized to test whether or not the patients in the organic group made scores significantly higher than patients in the other groups. For this particular analysis, the contradictory hypotheses concerning line quality were eliminated, and a score was obtained for each patient consisting of the number of indicators present out of a possible total of 11 (Nos. 1–11). These scores were converted to percentages, and an angular transformation was done in order to stabilize the variance. The analysis-of-variance design was then applied to the transformed scores of the 75 patients distributed into the three diagnostic groups.

Results and Discussion

From the results of the analysis of variance cited in Table 1, which yielded nonsignificant results, it is evident that the organic patient does not characteristically draw larger figures than other patients.

Tables 2 and 3 show that 5 of the 14 indicators are indeed more characteristic of organic drawings than of drawings of either

schizophrenics or neurotics. An additional indicator, emptiness of figures, showed a significant difference between organic and neurotic productions, but not between organic and schizophrenic drawings. The remaining indicators occurred with approximately equal frequency in all groups, with any discrepancies attributable to chance factors. When schizophrenics were compared with neurotics, none of the indicators differentiated between the groups to a statistically significant degree.

It may be noted here that Vernier's sketchy or broken line indicator was present more often in the organic group than was Machover's simple, heavy line. However, neither of these indicators appeared to a greater extent in the organic group than they did in the other groups.

Table 3

Drawing Indicators Showing Significant Differences Between Groups of Organic and Schizophrenic Patients of Twenty-five Each

Indicator	Frequency of occurrence: Organic	Schizophrenic	Chi square	*p*
Weak synthesis	15	5	6.75	.01
Parts misplaced	15	5	6.75	.01
Shrunken arms and legs	14	0	16.77	.001
Parts other than head and extremities distorted	15	6	5.26	.05
Petal or scribbled fingers	13	1	12.00	.001

Results of a final analysis are reported in Table 4 and indicate the highly significant degree to which organic patients surpass both schizophrenics and neurotics in the number of indicators present in their individual drawings.

Since the analysis-of-variance design applied to the transformed scores demonstrated a disparity among the three groups significant at the .001 level, the groups were next considered two at a time. The *t* test was applied using the mean square value of the within-group source of variation (error term) as the pooled variance. Results of the comparisons are shown in Table 5. It can be seen that the scores of the organic patients differed from both those of the neurotics and schizophrenics at the .001 level of confidence, while the difference between scores of the schizophrenic and neurotic patients was not statistically significant.

Table 4

Analysis of Variance of Transformed Drawing Scores of Organic, Schizophrenic, and Neurotic Patients

Source of variation	*df*	Mean square	*F*
Between groups	2	1612.19	13.65*
Within groups	72	118.15	
Total	74		

* Significant at the .001 level of probability.

Although these indicators distinguish between groups, it remains to be determined how effectively they can be utilized in individual differential diagnosis. Table 6 reveals that as many as 7 out of 11 indicators must be present in order to establish with certainty that the correct diagnosis is an organic one, rather than neurosis or schizophrenia. With such stringent requirements, one would miss nearly three-quarters of all organic cases! Reducing the number of indicators to 5 would still result in 40 per cent false negatives and also in obtaining 4 per cent and 16 per cent false positives from among the neurotic and schizophrenic groups, respectively. Since these predictions are based on the same cases as were used for the analysis of the indicators, it is likely that a greater number of false diagnoses would occur on cross validation with a new sample. The evidence clearly points to the limitations of a naive reliance on a purely "sign" approach to the

Table 5

Separate Group Comparisons of Drawing Scores of Organic, Schizophrenic, and Neurotic Groups

Comparison	*t*	*p*
Organics with neurotics	4.66	.001
Organics with schizophrenics	4.16	.001
Neurotics with schizophrenics	0.49	not significant

Table 6

Cumulative Distribution of Drawing Scores of Organic, Schizophrenic, and Neurotic Patients

Drawing score*	Organics		Schizophrenics		Neurotics	
	Number of patients	% of patients	Number of patients	% of patients	Number of patients	% of patients
1 or more	24	96	24	96	24	96
2 or more	24	96	20	80	19	76
3 or more	21	84	15	60	14	56
4 or more	19	76	7	28	6	24
5 or more	15	60	4	16	1	4
6 or more	12	48	1	4	1	4
7 or more	7	28	0	0	0	0
8 or more	4	16	0	0	0	0
9 or more	1	4	0	0	0	0

* Drawing score equals number of indicators recorded as present out of maximum of eleven.

diagnostic task. Nevertheless, the sophisticated clinician may well utilize the more discriminating indicators as a basis for further investigating the possibility of organic pathology in a patient.

Summary

This study attempted to evaluate a number of drawing indicators cited in the literature as characteristic of organic patients. It was found that five of these indicators were more prevalent among a group of organic patients than among schizophrenic or neurotic groups. These were: weak synthesis; parts misplaced; shrunken arms and legs; parts, other than head and extremities, distorted; and petal or scribbled fingers. The remaining ten indicators did not differentiate between the three groups.

When the number of indicators present was scored for the patient's individual drawings it was found that the organics had significantly higher scores, but the groups overlapped to a considerable degree, limiting the usefulness of this approach alone for individual differential diagnosis.

Received February 24, 1956.

References

1. Machover, Karen. *Personality projection in the drawing of the human figure.* Springfield, Ill.: Charles C Thomas, 1949.
2. Michal-Smith, H. The identification of pathological cerebral function through the H-T-P technique. *J. clin. Psychol.*, 1953, **9**, 293–295.
3. Vernier, Claire M. *Projective test productions: I. Projective drawings.* New York: Grune & Stratton, 1952.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Certain Formal Characteristics of Human Figure Drawings by Institutionalized Aged and by Normal Children[1]

Martin Lakin

University of Illinois, College of Medicine
Orthodox Jewish Home for the Aged, Drexel Home for the Aged
Chicago, Illinois

Drawings of the human figure have been fairly widely used and studied as a psychological diagnostic and personality assessment technique. They have been used as bases for inferences about intellectual or maturity level, conflict areas, and self concept or "body image." The present study is an investigation of the extent to which figure drawings reflect self concepts or body images which may be assumed to be characteristic of disparate levels of age and social status. The specific question under consideration is whether figure drawing is a technique which is sensitive to *changes* in body image occurring from childhood to old age.

That changes in body image might be anticipated in individuals who experience the physiological vicissitudes associated with aging is suggested by Schilder; "When we build up a persistent and coherent image of the body we build it up out of our emotional mood which is based upon biological tendencies . . . the changes in the biological unit reflect themselves in the more plastic image of the body" (4).

Several studies have found that the drawings of institutionalized aged persons differ from the drawings of certain groups of normal adults. According to one recent study, as compared with the productions of university graduate students, those of the aged were characterized by incompleteness, lack of proportionality, and lack of integration (2). Machover (3) has observed that senile individuals draw small, puny figures, use relatively little area of the paper for the drawing and center the figure poorly upon it.

These observations are interesting in view of current theory with respect to the technique. In clinical work, gross departures from "normative" performances in terms of such formal variables as area, size, and centeredness are frequently interpreted as suggestive evidence of deviant intrapersonal states. Area utilized, for example, is considered to reflect self-assessment in relation to the environment. Height is similarly interpreted as an indication of feelings of relative self-significance. Centeredness is often considered to be affected by the subject's general security level and by his orientation toward reality.

Gross differences in self conceptualization and, presumably, in body image would be anticipated in a comparative study of the institutionalized aged and normal children close to the threshold of adolescence. In one way their life situations are somewhat similar: both groups are explicitly dependent upon protecting and nurturant environmental forces for gratification of primary needs. However, as contrasted with the vigor, growth, increasing self-awareness, and recognition of potential among the children, the psychological decrements which accompany the physiological changes in the aged are a matter of popular concern. Loss of self-esteem, lessened vitality, and the sense of impending further decline seem to underlie the frequently observed ex-

[1] Thanks are due Drs. Marion Wieman and Alan K. Rosenwald of the University of Illinois, College of Medicine for their advice and cooperation in this study.

Table 1

Identifying Data for Subjects

Mean	Age	Education (yrs. of formal schooling)	*N*	
			Males	Females
Institutionalized aged subjects ($N = 24$)	78.5	3*	13	11
Normal children ($N = 25$)	8.7	3	10	15

* Estimated because of discrepancies in education of aged who were all foreign born.

pressions of futility and hopelessness among the institutionalized aged. The question is whether these differences are reflected in the figure drawings of the two groups.

Such formal variables as area, height, and relative centeredness would appear to be excellent points of comparison between the two contrasting groups. They are elements which are relatively free from the influences of previous education and artistic training. They are much less subject to these influences and to conscious control than such features as integration of the figure, facial expression, and proportionality. Also, with respect to these latter elements, the influence of peripheral, sensorimotor impairments would have to be taken into account. In the present study of formal characteristics peripheral impairment would seem to be of minimal significance. If the formal characteristics, area, height, and centeredness do reflect self-conceptual or body-image dispositions, then it would be hypothesized that differential performance in terms of them would distinguish between the drawings of institutionalized aged and those of normal children.

Method

Twenty-five third grade children from a public school and twenty-four elderly residents of a home for the aged were the subjects. All were administered the Figure Drawing Test (3) and were provided with similar materials, medium soft pencils and plain 8½ × 11 sheets of paper. The aged subjects were given the test in individual examination sessions. The children were tested as a group. Instructions were limited to the simple directions: "Draw a person; you may make it any way you wish, but draw a whole person, not just the face."

Table 1 describes the personal characteristics of the groups. The children were a typical class of third graders. Those markedly high or low in previously administered Stanford Binet scores (above IQ 110 or below IQ 90) were not included in the present sample. The institutionalized aged were also a fairly "typical" group, in terms of the home's standards. All were ambulatory and in "good" contact. Markedly deteriorated individuals and those considered to be grossly impaired intellectually as evaluated on the basis of the Bender Gestalt Test and psychiatric interview, were not included in the study.

Comparison of the aforementioned formal characteristics of the drawings were undertaken as follows:

Area. The area of each drawing was precisely measured by means of a geographer's planimeter. The instrument yields results in terms of square inches. The areas of the two groups, treated as scores, were then compared. Because of the somewhat skewed distribution of the data the nonparametric median test was employed.

Figure height. The distance from figure top to figure bottom was measured with an ordinary ruler. Heights to the nearest inch were treated as scores. The scores of the groups were then compared by means of the median test.

Centeredness. A grid of concentric circles, ¼ inch apart, was constructed. Each circle was numbered and regarded as a unit. By placing the grid-template over the sheet with the drawing, it was possible to measure the distance from the center of sheet to the center of gravity (the midpoint between the figure's extremities) of the figure. Measured distances from the center in units were termed scores and were compared by means of the median test.

Results

Table 2 shows the differences in the performances of the two groups of subjects.

On the measure of area the group of institutionalized aged and that of normal children had medians of 3.0 square inches and 6.8 square inches respectively. On the second measure, that of distance from the center in ¼ inch units, the medians were 11.0 and 6.0.

Medians of the group with respect to the measure of height were 4.0 inches and 6.0 inches respectively.

To provide some indication of the variability the number exceeding the combined median is given for both groups. For the measure of area the numbers are 6 and 18 respectively. For distance from center they are 17 and 8, and for height they are 6 and 19.

Discussion and Conclusions

The results confirm the hypothesis that formal characteristics of the human figure drawings differentiate two groups widely differing in age and social status, and assumed to be widely differing in self concept and body image.

The direction of the difference, the more constricted, shorter, and less adequately centered productions of the institutionalized aged, as compared with those of the children is congruent with the assumed differences in self-attitudes among these groups. What is significant, in terms of the present study, is that the feelings of self-devaluation and a "shrinking" body image should find graphic expression in the execution of formal features of the drawings. On the other hand, the possibility that certain organic factors in the aged result in constrictions of the figure is counterindicated by nonconstricted Bender-Gestalt drawings among this group.

The findings seem to give additional evidence of the validity of the assumption that formal aspects of figure drawings are related to the central variables of self conceptualization and body image. In light of the clinical assumption about departures from "normative" performance it is interesting to note that the children seem to perform more similarly to normal adults than do the aged subjects. Levy (1) reported an average figure drawing height of 7.0 inches for normal adults. The median figures for the aged and child subjects of the present study are 4.0 inches and 6.0 inches respectively.

The findings suggest the application of similar methodology in further investigation in this area. Comparison of drawings by institutionalized aged and by noninstitutionalized aged persons might be undertaken to study the question of whether changes in self concept and body image are due to psychological processes attendant upon aging per se or whether they are related to such factors as relative independence of functioning and the effects of permanent custodial-protective care.

Table 2

Differences in Area Size, Height, and Distance from Center of Figure Drawings of Institutionalized Aged Persons and Normal Children

	Group		
	Institutionalized aged	Normal children	Median test
Area (sq. inches)			
Median	3.0	6.8	
Number exceeding combined median	6	18	10.8**
Distance from center ($\frac{1}{4}$" units)			
Median	11.0	6.0	
Number exceeding combined median	17	8	7.4*
Height (to nearest inch)			
Median	4.0	6.0	
Number exceeding combined median	6	19	12.7**

* Significant at the .01 level.
** Significant at the .001 level.

Summary

On the basis of reported observations and theory regarding human figure drawings and their relationships with self concept and body image it was considered that such formal characteristics of the drawings as area, height, and centeredness would reflect aspects of change in self concept and body image. The relative freedom of these variables from background influences and from the influences of peripheral impairments was suggested. It was hypothesized that differences in self conceptualization between institutionalized aged and preadolescent children would be reflected in differences in these formal aspects of the drawings. The drawings of 24 institutionalized elderly persons were compared with those of 25 third-grade school children. The re-

sults confirmed the hypothesis. More constricted, shorter, and less adequately centered figures were produced by the aged subjects. The findings were viewed as support for the validity of the view that the central variables of self concept and body image are reflected in drawings of the human figure. The application of this method in further research in this area was suggested.

Received July 11, 1956.
Early Publication.

References

1. Levy, S. Figure drawing as a projective test. In L. E. Abt & L. Bellak (Eds.), *Projective psychology*. New York: Knopf, 1950. Pp. 257–297.
2. Lorge, I., Tuckman, J., & Dunn, M. B. Human figure drawing by younger and older adults. *Amer. Psychologist*, 1954, **9**, 420–421. (Abstract)
3. Machover, Karen. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Charles C Thomas, 1948.
4. Schilder, P. *The image and appearance of the human body*. London: Kegan Paul, 1935.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Cue-Counting: A Measure of Anxiety in Interviews

Andrew S. Dibner
Veterans Administration Hospital, Dearborn, Michigan [1]

Previous studies relating formal aspects of speech to anxiety (2, 3, 4, 5, 6) almost entirely concern themselves with the reflection in speech of anxiety as a more or less permanent concept of personality, i.e., the anxiety of the "anxious" or "disturbed" individual. There have been no speech measures reported which were developed to record anxiety as a transient, situational phenomenon, although transient anxiety is a relevant issue when one studies the psychotherapeutic process. Such speech measures were developed by the author in a study relating patients' anxiety in an interview to the ambiguity of the interviewer's behavior (1). It was desired to have speech indices as well as other indices of anxiety. The indices developed standardize the measurement of clinical "cues" of anxiety appearing in speech.

Method

Subjects

Subjects were 39 patients hospitalized on the NP Service of a VA hospital. Diagnoses were all of psychoneurosis, character disorder, or psychosomatic reaction. Subjects ranged in age from 25 to 54, had a mean IQ of 105 (range 85–115) and mean education at the 10.5 grade level.

Speech Material

Transcriptions of tape recordings of initial psychological interviews with the subjects provided the speech material from which the measures were developed. Verbatim typing of the transcripts was done by trained typists following instructions which standardized indications of pauses, sighing, laughing, and changes in loudness of voice.

[1] Now at Clark University, Worcester, Mass.

Procedure

Careful study of transcripts of 8 interviews gathered in a pilot study revealed 11 speech characteristics which we believed were scorable and related to situational anxiety. These are listed and illustrated below.

UN—Unfinished Sentence. The thought is interrupted and never completed; another thought might be expressed, the old thought approached in a radically different way, or silence might occur. Examples:

1. "I went to live with my Dad and I had . . . (UN) I lived in (city) with him until I was 16."
2. "Oh, she uh . . she's a (R) . . (UN) I mean she always stayed at home."
3. "Well, that's the main . . . (silence) (UN)."

BR—Breaking in with a New Thought, generally by breaking into another sentence. Also if a new sentence introduces a radically different line of thought and there is no sign of a preceding pause. Examples:

1. "When I was young my mother . . (UN) I was (BR) . . uh . . ashamed of myself, very much ashamed."
2. "And when I had this here uh . . depression . . (INT) uh, I know what I'm saying right now (BR) . . . uh, I had quite a jag."

INT—Interrupted Sentence. The thought is interrupted within the sentence but is eventually completed in much the same manner as it was begun. Examples:

1. (See example 2 under BR above).
2. "I was going . . (INT) see, nobody knew it, but I decided I was going home."

R—Repeating Words or Phrases. When full words or phrases (either parts or full) are repeated close in sequence within a sentence. Also included are modifications of words or phrases. Examples:

1. "I don't have . . I don't have (R) as much gall as some people."
2. "I worked for two years in a . . an (R) office when I was 16 . . no, 17 (R)."

ST—Stuttering. When repeating parts of words or when leaving a word unfinished. Examples:

1. "I ne . . never (ST) thought of it like that."
2. "He even sta . . . (ST) he even stayed (R) away from me."

IDK—I Don't Know. When the patient says "I don't know," not in answer to a direct question, but as in resignation, disgust, or despair. Examples:

1. "So that's the way it happened. I don't know . . . (IDK) things were all mixed up."
2. ". . . and the pain has been there since then. (Pause) I don't know (IDK)."

SI—Sighing or Deep Breath.

L—Laughing. Any kind of laugh or chuckle.

V—Voice Change. Change in volume, becoming louder or lower.

?—Questioning the Interviewer. When the patient asks the interviewer to repeat what he said, not for elaboration but because he was not heard or understood. Examples:

1. P: "What?" (?)
2. P: "Pardon me?" (?)

BL—Blocking. When there is groping for the proper expression, indicating unusual hesitation. Examples:

1. "So she . . uh . . well . . . uh . . (BL) told me I shouldn't come back."
2. "So I . . so I (R) . . (snaps fingers) (BL) I (R) told her I wouldn't anyway."

To estimate reliability of scoring these cues, two raters trained together by reviewing and discussing the scoring of eight 10-minute samples of interviews. Then they separately scored cues on 10 interviews and their percentage of agreement on each cue was computed by dividing the number of agreements by the total of agreements plus disagreements. Agreement means both raters note the same cue in the same spot on a line of speech. Disagreement occurs when either rater notes a cue and the other calls it some other cue or leaves that spot unmarked.

Table 1

Percentage of Agreement of Two Raters on Each Cue

Cue	Agree/agree plus disagree	% Agreement
UN	82/134	61
BR	5/26	19
INT	24/72	33
R	143/213	66
ST	16/17	94
IDK	9/19	47
SI	14/14	100
L	22/22	100
V	41/41	100
?	4/4	100
BL	2/2	100

Ten-minute samples of the remaining 29 typescripts of interviews were then cue-counted and a Cue Score, expressed as a percentage, was computed for each interview by dividing the total number of times the cue was scored in the interview by the total number of typed lines of patient speech. To analyze the interrelationships of the cues, a contingency coefficient between each distribution of cue scores was then computed. The intercorrelations indicated the groups of cues which we called Cue Scale I and Cue Scale II.

Group cue scores for Cue Scale I and Cue Scale II were obtained by converting the cue score distributions over all interviews to percentile scores, then for each interview adding the percentile cue scores within each scale. Because of its negative correlation with V and SI, the L distribution was made negative before combining it with them.

Results

Table 1 presents the raters' agreement on scoring each cue. The perfect rater agreement on SI, L, V, ?, and BL results from the fact that they are clearly marked on the typescript by the transcribing secretary. UN, R, and especially ST have respectable per cent agreements, but it was obvious that if BR, INT, and IDK were to be used, they would have to be grouped with the more reliable cues.

Table 2 presents the intercorrelations of the cue scores. The two cues ? and BL were

Table 2

Intercorrelations of Cues
Contingency Coefficients Corrected for Class Index
(39 Interviews)

Cue	BR	INT	R	ST	IDK	SI	L	V
UN	.42	.20	−.03	.20	.56	−.20	−.13	.13
BR	—	.03	.13	.20	.03	.13	−.28	.13
INT	—	—	.36	.56	.60	.03	.13	−.28
R	—	—	—	.63	.13	.28	−.13	−.20
ST	—	—	—	—	.20	.24	−.05	.11
IDK	—	—	—	—	—	.28	−.28	−.03
SI	—	—	—	—	—	—	−.28	.69
L	—	—	—	—	—	—	—	−.28

omitted from consideration because they occurred rarely in the data and provided distributions with very little spread.

By inspection of Table 2, the cues were grouped into patterns containing the highest possible intercorrelations. The following groupings were prominent: BR-UN-IDK-INT-R-ST and L-V-SI. These groups we call respectively Cue Scales I and II. With the exception of IDK, the cues of Scale I are all indicators of speech disruption. Scale II is composed of cues related to nonverbal voice characteristics. The reliability of rating of the cues grouped in Scale I is 65 per cent and of those grouped in Scale II is 100 per cent. Cue Scales I and II have a Kendall's rank correlation (Tau) of $-.03$ ($p = .81$). They are unrelated and used independently as anxiety measures.

Discussion

In the author's study referred to earlier (1), the Cue Scales and three other indices of anxiety (skin conductance curves, clinicians' global ratings, and patients' self-reports) were used to measure patient anxiety in reaction to structured or ambiguous interviewing technique. The Cue Scales showed significant relationships to degree of ambiguity, just as did the other measures. The anxiety measures showed no significant correlation with each other, except for the correlations (.32 and .26) of the two Cue Scales with anxiety ratings of transcripts by clinical judges. This correlation indicates only that both sets of speech cues (uncorrelated with each other) enter into a judge's estimate of anxiety when reading transcripts. An analysis of the findings of the study supports the hypothesis that anxiety is not a unidimensional phenomenon but can be evidenced by activation of various physiological or behavioral systems in different persons. Thus a person who shows speech disruption in a situation of uncertainty need not show concomitant GSR changes. Is GSR a better indicator of anxiety than speech disruption? At this point it seems safe to say only that they are both methods by which we can identify anxiety in order to learn more about it. The speech disturbances measured by the Cue Scales are clearly related to situational anxiety and not to anxiety as a personality trait, for the patient groups showed significant speech differences in reaction to the varied interviewer technique although they were randomly selected and equated for diagnosis and several other pertinent variables.

The Speech Scales represent speech variables indicative of "emotionally involved" and disorganized verbal behavior. Some of the variables seem to be direct instances of disorganization (e.g., incomplete sentence or breaking in of new thoughts) and others as defenses against or attempts to control disorganization (e.g., repetition or modification of words and phrases). The lack of correlation (Tau = .01 and .07) of the two scales with the patient's own report of anxiety leads us to feel that the cues of disorganization, involvement, and defense, measured by the scale, are anxiety indicators not subject to the person's awareness. Further research should be done to explore other speech cues of anxiety, to evaluate the consistency of appearance of the cues under varying conditions of anxiety, and to investigate the relationship of the cues to individual speech habits. The use of the scales as tools to measure anxiety related to such therapeutic phenomena as interpretation, reassurance, or transference, as well as for evaluation of therapeutic progress should be explored.

Summary

The derivation of two measures of transient or situational anxiety as revealed in speech has been described. The two uncorrelated measures are based upon a count of speech cues reflecting disorganization, emotional involvement, and defensiveness. The use of the scales in a study of anxiety in psychological interviews was described, and some implications of the scales as anxiety measures were discussed.

Received March 12, 1956.

References

1. Dibner, A. S. The relationship between ambiguity and anxiety in a clinical interview. Unpublished doctor's dissertation, Univer. of Michigan, 1953.
2. Dollard, J., & Mowrer, O. H. A method of measuring tension in written documents. *J. abnorm. soc. Psychol.*, 1947, **42**, 3–32.
3. Kauffman, P. E., & Raimy, V. C. Two methods of assessing therapeutic progress. *J. abnorm. soc. Psychol.*, 1949, **44**, 379–385.
4. Lasswell, H. D. Verbal references and physiological changes during the psychoanalytic interview: a preliminary communication. *Psychoanal. Rev.*, 1935, **22**, 10–24.
5. Sanford, F. H. Speech and personality. *Psychol. Bull.*, 1942, **39**, 811–845.
6. Zimmerman, J. Modification of the DRQ as a measure of progress in counseling. Unpublished master's thesis, Univer. of Chicago, 1950.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Test Anxiety and Intelligence Test Performance

I. Zweibelson
Milford, Connecticut, Public Schools

Relatively few persons have escaped the worry which is usually associated with the group tests which are used extensively in our culture. It is also true that there have been relatively few studies of the influence of fear, worry, and anxiety on group test performance. Studies (3, 7) of the worries of elementary school children have revealed that they worried most about test failure. It has also been shown (8) that college students with high anxiety performed noticeably more poorly on group intelligence tests than students with low anxiety.

The Problem

The present study [1] concerns the relation between children's test worry or anxiety and their scores on group tests of mental ability which differed in content, format, and method of administration. The following question was posed: does performance in mental ability test situation x, y, or z differ in relation to the relative amount of test anxiety indicated by the subject? Table 1 briefly summarizes the important differences of the three mental ability tests used in the study: the Otis-Alpha, the Davis-Eells, and the Otis-Beta. An index of pupil test anxiety was secured by the investigator who read a list of 43 questions [2] to 12 classes of fifth grade elementary school children. Each child responded on an answer sheet by circling a *"Yes"* or *"No"* for each question. The questions dealt specifically with fears, worries, and feelings in test and testlike situations.

[1] Adapted and condensed from a dissertation submitted in partial fulfillment of the requirements for the Ed.D. degree at Teachers College, Columbia University, in 1955. Part of the data were collected in connection with the USPHS Project M-712 under the direction of Professor Seymour B. Sarason of Yale University.

[2] The questionnaire used was developed under the direction of Professor Sarason of Yale.

Hypotheses

1. Given three tests of mental ability, the Davis-Eells, the Otis-Alpha, and the Otis-Beta, performance on the Davis-Eells and the Otis-Alpha tests should be less related to test anxiety than in the case of the Otis-Beta.
2. In the case of the Otis-Beta and the Davis-Eells, the "High-Anxious" should do better on the mental ability test taken second than on the one taken first, regardless of the test, i.e., the order in which the tests are taken will have an effect on the scores of the "High Anxious."

Design

Five of nine elementary schools in Milford, Connecticut, were selected as schools which served town areas thought to be representative of pupils in all of the town schools. All of the fourth grade children in the selected schools were given three mental ability tests (1, 5, 6) and the Test Anxiety Questionnaire. The Otis-Alpha and the Questionnaire were administered in the ninth month of the fourth grade school year. The Otis-Beta and the Davis-Eells were administered during the first month of the fifth grade school year. Half of the pupils were given the Otis-Beta first and half the Davis-Eells first—the total N was 258. There was approximately a week between the administration of the first and the second tests.

Results

The correlations among the three mental ability tests were:

Table 1

Major Differences of Three Group Mental Ability Tests

Factor	Difference		
	Otis-Alpha	Davis-Eells	Otis-Beta
Test Items	variety of problems include geometric, pictorial, and language ability items	real-life problems, personal oriented	school achievement
Style	semiacademic	semihumorous	academic
Method of Administration	semiformal test atmosphere	game atmosphere stressed	formal test atmosphere
Instructions	little provision for praise and reassurance	substantial provision for praise and reassurance	little provision for praise and reassurance
Format	free of reading	free of reading	reading required
	small, clear pictures	large, clear pictures	print
	picture or symbol marked	large space for mark next to test item	small space on separate sheet for mark
	booklet form	booklet form	3 sheets fold like accordian
Provision for Practice	short period preceding each of two sittings	substantial periods preceding several sections	short period preceding the one sitting
Time Limit	40–50 minutes	100–120 minutes	30 minutes
Speed Factor	emphasis tends to be on power although some don't have time to try all of the items	emphasis upon power, each pupil could attempt every item	speed important, no pupil could attempt every item

Otis-Alpha vs. Davis-Eells	.59
Otis-Alpha vs. Otis-Beta	.55
Davis-Eells vs. Otis-Beta	.49

These correlations are significant at the one per cent level of probability.

The correlations among the tests and test-anxiety scores were:

Otis-Alpha vs. Text Anxiety	− .28
Otis-Beta vs. Test Anxiety	− .24
Davis-Eells vs. Test Anxiety	− .14

These correlations are significant at the five per cent level. The significance of the differences between the Test Anxiety and the intelligence test correlations was found by the use of a formula developed by Hotelling and given by Walker and Lev (9, p. 257). In this formula the correlations among the tests are taken into account. The difference between the Test Anxiety vs. Otis-Beta r of − .24 and the Test Anxiety vs. Davis-Eells r of − .14 was not found to be significant, although the results are in the predicted direction ($p = .10$). The difference between the Test Anxiety vs. Otis-Alpha r of − .28 and the Test Anxiety vs. Davis-Eells r of − .14 was found to be significant at the five per cent level of probability. It appears that the first hypothesis receives support from these results although it is most clear in terms of the Otis-Alpha vs. Davis-Eells data.

Table 2 shows that the "low" test-anxious make higher mental ability indices than the "high" test-anxious regardless of the order of test administration and regardless of the test

instrument. There is also a tendency for the "High Anxious" to do better on the test taken second than on the test taken first. However, an analysis of variance of the data reported in Table 2 shows no significant variation in the size of the difference between the Davis-Eells and Otis-Beta indices because of: (*a*) the effect of the extent of test anxiety; or (*b*) the effect of the order in which tests were taken; or (*c*) the effect of the interaction of these variables. It should be pointed out, however, that the order of test administration was the greatest source of variation of the several sources investigated. The data in Table 2 clearly do not allow one to accept Hypothesis 2, although the fact that the trends are in the predicted direction suggests that consideration of other factors in future studies may give more positive results. It might be pointed out in this connection that in studies of test anxiety with college students (2, 4) and children there is evidence that the "High Anxious" subjects tend to improve rather markedly in the course of learning.

Table 2

Mean and Variance of Mental Ability Test Scores of High and Low Test-Anxious Pupils Within Order of Test Administration
(N = 33 in each group)

	Pupil test anxiety			
	Lower 31%		Upper 31%	
Order of administration	$\bar{X}$	s^2	$\bar{X}$	s^2
Otis-Beta Administered First				
Otis-Beta Score	5.5	3.9	4.0	2.2
Davis-Eells Score	5.4	3.8	4.4	2.4
Davis-Eells Administered First				
Otis-Beta Score	5.0	3.1	3.9	4.2
Davis-Eells Score	4.7	3.1	3.4	3.5

Discussion

The results of the present study support the position that test anxiety may be related to and effect group test results. There is also some support for the Davis and Eells contention that their test is less influenced by anxiety than other conventional tests. Evaluation-connected experiences which involve failure or threat of failure may serve to engender personal reactions which interfere with test performance. The standardized group test situation may have threatening factors inadvertently attached to the experience by the test-maker or the examiner.

The anxiety-provoking role of the examiner has not been studied in this present investigation. The different examiners (teachers and psychologists) in this study may have influenced the results by provoking different anxious reactions in the children. It seems important to draw attention to such possibilities and to suggest that the examiner's role in effecting group test performance might warrant investigation.

Received April 9, 1956.

References

1. Davis, A., & Eells, K. *Davis-Eells test of general intelligence or problem-solving ability.* Yonkers, N. Y.: World Book Co., 1953.
2. Gordon, E. A. *Self esteem, need for achievement, test anxiety and performance.* Unpublished doctor's dissertation, Yale Univer., 1955.
3. Jersild, A. T., Goldman, B., & Loftus, J. J. A comparative study of worries of children in two school situations. *J. exp. Educ.*, 1941, **9**, 323–326.
4. Mandler, G., & Sarason, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47**, 166–173.
5. Otis, A. S. *Otis Quick-Scoring Mental Ability Tests. Alpha Test, Form B.* Yonkers, N. Y.: World Book Co., 1936.
6. Otis, A. S. *Otis Quick-Scoring Mental Ability Tests. Beta Test, Form Cm.* Yonkers, N. Y.: World Book Co., 1939.
7. Pintner, R., & Lev, J. Worries of school children. *J. genet. Psychol.*, 1940, **56**, 67–76.
8. Sarason, S. B., & Mandler, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 810–817.
9. Walker, H. M., & Lev, J. *Statistical inference.* New York: Holt, 1953.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The Relation of the Children's Manifest Anxiety Scale to the Concept of Anxiety as Used in the Clinic[1]

Robert D. Wirt and William E. Broen, Jr.[2]
University of Minnesota

Considerable research has been published in recent years indicating relationships between a construct measured by the Taylor Manifest Anxiety Scale (MAS) and other variables mainly within the scope of learning theory. Taylor recently has reviewed this work (2). More recently Castaneda, McCandless, and Palmero have adapted the MAS for use with children (CMAS) and have reported research indicating that its relationship to other variables approximates that of the adult form of the scale (1).

Because of the importance of the anxiety construct to clinical work, we investigated the relationship between the CMAS and the concept of anxiety as used in the clinic. Using the CMAS we tested 505 children in the 4th, 5th, and 6th grades in the Minneapolis Public Schools and 34 children from the same grades who had been referred to its Child Study Department. Predicting from the clinical concept of anxiety as we understand it, we hypothesized that the clinic children would score higher on the CMAS than the others. Analysis of variance failed to support this expectation. We also predicted there would be a positive relationship between anxiety as rated by psychologists after diagnostic study or as rated by teachers and scores on the CMAS (not verified); that intelligence would be inversely related to CMAS scores (not verified); and that there would be an inverse relationship between CMAS and reading achievement scores with intelligence partialled out (significant at the .04 level).

As a result of this research we feel that the CMAS (and the MAS insofar as it is comparable) although undoubtedly a valuable and predictive instrument is measuring something different than is denoted by the clinical concept of anxiety. We find no contradiction between this conclusion and work cited by Taylor (2) which shows a relationship between the MAS and other clinically used variables. We believe the common factor underlying the relationships to be one of "willingness to say deviant things about the self" or "test-taking attitude" which factor is not necessarily highly related to anxiety as the concept is clinically used. Studies showing significant relationships have emphasized *manifest* anxiety, as, of course, do the names of the scales.

As has been pointed out by others, concepts such as anxiety need much more careful denotation than they have yet been given. We are presently engaged in further research on the construct of anxiety with the hope of giving it greater specificity.

Brief Report.
Received July 30, 1956.

[1] An extended report of this study may be obtained without charge from Robert D. Wirt, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota, or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. 5038 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

[2] Now at the University of Kansas.

References

1. Castaneda, A., McCandless, B. R., & Palermo, D. C. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317–326.
2. Taylor, Janet A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, 53, 303–320.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

The "Alienation Syndrome" and Adler's Concept of "Distance"

H. L. Ansbacher
University of Vermont

In a paper published in this Journal and entitled "Alienation, social apperception, and ego structure," Anthony Davids (2) reported an Affect Questionnaire which he constructed in collaboration with Henry A. Murray. The questionnaire measures eight critical dispositions of personality: sociocentricity, optimism, and trust (three "positive" or socially desirable dispositions); and egocentricity, pessimism, distrust, anxiety, and resentment (five "negative" or socially undesirable dispositions). Statistical analysis of results obtained with this questionnaire from Harvard students led to the formulation of a syndrome termed "alienation," which is composed of the five negative dispositions.

The purpose of the present note is to show that the concept of "alienation" bears close resemblance to Adler's concept of "distance." Etymologically, the two terms have in common the idea of isolation from the group (like the sociologists' term "anomie"), the first denoting the feeling of being a stranger, the second, the attitude of standing apart from the rest.

Adler's "distance" manifests itself in a fourfold modus of, figuratively, moving backward, standing still, hesitating, or constructing obstacles (1, pp. 273–276). All kinds of anxiety are seen as forms of distance behavior. To anxiety so described, Adler subordinates a number of specific feelings and attitudes which bear close similarity to Davids' negative dispositions. In order to demonstrate this, we shall insert in the following quotation from Adler the corresponding descriptive terms from Davids.

> Once a person has acquired the attitude of running away from the difficulties of life, this attitude may be greatly strengthened and safeguarded by the addition of anxiety (Davids' *anxiety*). . . . It embitters (Davids' *resentment*) his life to a marked degree, renders him quite unsuited for making contacts, and thus also for laying the basis for a peaceful life and for fruitful accomplishments. . . . We shall meet again the well-known type who feels himself forced by necessity to think more of himself and consequently has little left for his fellow man (Davids' *egocentricity*). As though in confirmation of this view one often finds that such people are inclined to think of the past or of death, which has about the same effect. . . . Or they emphasize that all is vanity, after all, that life is short, and one cannot know what will happen (Davids' *pessimism*)" (1, pp. 276–277).

This quotation permitted the insertion of all of Davids' negative dispositions except *distrust*, which is found in Adler under the heading of Oversensitivity, one of the characteristics of the neurotically disposed individual. "Distrust in his own strength which awakens doubt about everything; distrust in others which is socially disturbing and disrupts every group—these at times characterize the oversensitive patient" (1, p. 290).

With these similarities, it is understandable that the further findings of Davids are in accordance with Adlerian theory. This is true especially for the finding "that individuals who view the world and their relationship to it in a negative (i.e., alienating) light have a weak ego structure (i.e., poor personal and social adjustment) as estimated by experienced clinical judgment" (2, p. 27). Adler had started with the observation that poor adjustment is characterized by the syndrome of distance which is the resultant of increased inferiority feelings and lack of social interest.

Davids developed the Affect Questionnaire to investigate relations between degree of alienation, social apperception, and ego structure, meaning ego strength vs. ego weakness. Not only by the formulation of the concept of alienation, but also by this purpose, is Davids' investigation in line with Adlerian theory. The relationship, according to Adler, would be: The maladjusted (alienated) individual is one who, due to increased inferiority feelings (ego weakness) and concomitant lack of social interest, apperceives the world as a hostile place. "Schema of apperception" is a term frequently employed by Adler. Specifically, the content of the schema of apperception of the maladjusted would be: They "feel and live as though they were in an enemy country. Fighting, hesitating, stopping, escaping, much more occupied with their own persons than with others, they are therefore selfish, inconsiderate, lacking in social interest, courage, and self-confidence because they fear defeat more than they desire success" (1, p. 118).

Altogether, the paper by Davids can well be taken as experimental support and extension of Adlerian theory.

Received February 23, 1956.

References

1. Ansbacher, H. L., & Ansbacher, Rowena R. (Eds.) *The individual psychology of Alfred Adler in selections from his writings.* New York: Basic Books, 1956.
2. Davids, A. Alienation, social apperception, and ego structure. *J. consult. Psychol.*, 1955, **19**, 21–27.

Journal of Consulting Psychology
Vol. 20, No. 6, 1956

Some Comments on Ansbacher's Note on the "Alienation Syndrome" and Adler's Concept of "Distance"

Anthony Davids
Brown University and Bradley Home

It is indeed gratifying when another investigator is able to make use of one's findings in building a bridge between empirical data and theory. This is especially true when the data were not gathered with a view toward supporting or refuting hypotheses derived from the particular theory with which they are eventually linked. The fact that the theory in question, that of Adlerian psychology, is not particularly new on the psychological scene, coupled with the fact that this theory has been submitted to very little experimental verification in controlled experiments, makes it seem all the more noteworthy that Ansbacher (2) was able to use our recent experimental evidence to support Adler's theoretical statements.

First, let me make a few comments on the dispositions which form the personality syndrome of alienation. On numerous occasions I have been asked why we chose to work with these particular eight dispositions of personality. That is, why did we include just these eight dispositions in our researches? Murray and I were first faced with the problem of choice of attributes to study when we embarked upon the task of constructing an auditory projective technique (5) several years ago. We wanted to include in this auditory test spoken statements and phrases indicative of certain critical dispositions of personality. By means of selective perception, apperception, accentuation, and retention of the spoken statements we hoped to secure a measure of the extent to which the perceiver possessed the various dispositions in his own personality. In our original pilot studies we experimented with many variables in addition to those which have now come to be regarded as forming an "alienation syndrome." For example, we employed other dispositions or attributes, such as self-pity, gratitude, activity, passivity, and several others.

On the basis of these pilot studies, we found that it was difficult to differentiate sharply among certain of the dispositions. That is, we found that both in writing statements to represent certain dispositions and in attempting to score subjects' responses, certain dispositions tended to overlap and some seemed not to be assessed with as much validity as others. Consequently, on the basis of practical considerations and preliminary studies, we decided to focus on the eight dispositions which are currently measured by both the auditory projective technique (5) and the Affect Questionnaire (4).

It should be pointed out, however, that deciding upon these dispositions was not merely an empirical or a theoretical development. These dispositions were proposed by Murray as being of critical importance based on his theory of personality dynamics and on his years of intensive investigation of normal and abnormal personalities. If one looks into Murray's contributions to the field of personality assessment (e.g., 7, 8, 9) it is evident that these eight dispositions, and several others, have long played a prominent role in his theorizing and his empirical investigations. Thus, the selection of these eight dispositions for study was based upon Murray's wisdom,

experience, and clinical knowledge about traits or dispositions that are of critical importance in the dynamics of personality.

Turning now to consideration of the Affect Questionnaire, it was constructed originally as an instrument against which to validate findings derived from the auditory projective test. This technique attempts to assess the relative strength of the eight dispositions in subjects' personalities by having them indicate their degree of agreement or disagreement with statements designed to represent the various dispositions. It was found that two experienced clinical psychologists could agree perfectly in their classification of the statements as indicative of the different dispositions. When the Affect Questionnaire (which consists of 10 statements representative of each of the eight dispositions, making a total of 80 items) was administered to groups of subjects, it was found that there were highly significant intercorrelations among the five negative dispositions, and that scores on these dispositions were negatively correlated with scores on each of the three positive dispositions. Thus, on the basis of these empirical findings (4), which were supported by unpublished findings obtained with additional groups of both normal and maladjusted subjects, we labeled the interrelated negative dispositions as forming an "alienation syndrome."

From the brief description of the selection of dispositions for study and the empirical findings which led to formulation of the syndrome, it is quite evident that the link with, or similarity to, Adler's concept of "distance" was not readily apparent to us, nor were our researches motivated by Adlerian theory. Thus, it seems quite striking that our empirical findings fit so neatly with Adler's theoretical formulations, and that these dispositions do, in fact, tend to coexist in the personality just as Adler believed on the basis of his theoretical convictions and his clinical experience.

Today much research attention is being devoted to submitting to experimental test many of the theoretical notions put forth by Freud on the basis of his clinical experience. However, this does not seem to be the case with Adlerian theory. It is, therefore, especially interesting to find that an experiment such as ours, which fits into the currently popular area of research termed "social perception" and which uses experimental techniques and statistical procedures, fits so nicely with hypotheses that would be generated by Adlerian theory. In other words, if we had taken as our point of departure the theoretical notions and statements advanced by Adler, and had designed our researches specifically to test propositions derived from his theory, we would have confirmed them in a well-controlled laboratory investigation employing modern statistical procedures. Since there is, undoubtedly, much of value in Adler's theory (3), and since many features or derivatives of Adlerian theory are put to use in present-day clinical practice (1, 6, 10), it seems of real value that Ansbacher was able to perceive the link between our seemingly unrelated experiment and Adler's theoretical concepts.

Received February 23, 1956.

References

1. Ansbacher, H. L. "Neo-Freudian" or "Neo-Adlerian"? Report on a survey conducted among members of the American Psychoanalytic Association. *Amer. Psychologist*, 1953, **8**, 165–166.
2. Ansbacher, H. L. The "alienation syndrome" and Adler's concept of "distance." *J. consult. Psychol.*, 1956, **20**, 483–484.
3. Ansbacher, H. L., & Ansbacher, Rowena R. (Eds.) *The individual psychology of Alfred Adler in selections from his writings*. New York: Basic Books, 1956.
4. Davids, A. Alienation, social apperception, and ego structure. *J. consult. Psychol.*, 1955, **19**, 21–27.
5. Davids, A., & Murray, H. A. Preliminary appraisal of an auditory projective technique for studying personality and cognition. *Amer. J. Orthopsychiat.*, 1955, **25**, 543–554.
6. Munroe, Ruth L. *Schools of psychoanalytic thought*. New York: Dryden, 1955.
7. Murray, H. A. *Explorations in personality*. New York: Oxford Univer. Press, 1938.
8. Murray, H. A. *Thematic Apperception Test Manual*. Cambridge: Harvard Univer. Press, 1943.
9. Murray, H. A., & Morgan, Christiana D. A clinical study of sentiments. *Genet. Psychol. Monogr.*, 1945, **32**, 3–149.
10. Thompson, Clara. *Psychoanalysis: evolution and development*. New York: Hermitage House, 1950.

PSYCHOLOGICAL TEST REVIEWS

Blum, Gerald S. *The Blacky Pictures.* New York: Psychological Corp., 1950.

Reviewed by

Samuel J. Beck

University of Chicago

A homely, obstreperous, and by now not unloved canine, who plays the title—and perhaps a hero's—role, this is Blacky of the Blacky test. Sired by psychoanalysis out of TAT he was created *ad hoc:* to test out psychoanalytic theory concerning psychosexual development, *Videlicet,* the title of Blum's monograph (2). It is a bold ingress into that most fluid area of psychology—depth psychology—which Blum undertakes, one more specific in purpose and design than those which projective tests have so far essayed. These, pioneers that they are, opening up a totally new field, have been engaged in attempting some broad estimate of the new territory, appraising its potential yield in terms of fresh knowledge concerning the human personality or, at the least, fresh ways of studying it. They have been treading warily, with some trembling at the shapes and sights unholy lurking in the scientific jungle that surrounds the regions of the human personality. However, as at other frontiers, the pioneers are followed by the more intensive cultivators. These latter can stake out, each his special acre, and so a method like that of the projective tests advances, by new experiment.

Blum has devised a test *ad hoc* for psychoanalytic theory by the ingenious technique which activates certain of the psychoanalytically formulated unconscious traits, at the same time that it provides the defenses enabling the patient to liberate and communicate these forbidden trends—the defenses of identification and displacing. I am not using any of the space allotted to me for this review in describing the test. He who has not yet made the acquaintance of the very human Blacky is referred to his creator's monograph (2). In carrying on his experiment, Blum used, as his point of departure, thirteen "dimensions" in psychoanalytic theory. Anyone even superficially read in this field will recognize the terms. His results, in University of California students, showed statistically significant differences between the sexes, statistically significant relationships between dimensions, and evidence in the writings of Freud and of Fernichel (he quotes the text from both) showing an impressive consistency between Blum's significant findings and psychoanalytic theory. He arrives at the cautious conclusion that "some aspects of psychoanalytic theory have demonstrable validity."

In the few years since he published, his test and findings have been subjected to fairly extensive research, not surprising what with the dominance of both psychoanalysis and projective instruments on the current psychological scene. Within a year, Seward criticized the exclusion of statistically insignificant but relevant data in Blum's report, and also pointed to some likely logical flaws (11). Michal-Smith, Hammer, and Spitz found their Blacky test analysis consistent with indication of Oedipal intensity in a nine-year-old boy (8). Paranoids were studied by Aronson (1) and he reported partial confirmation of Blum's results. Other investigations that have come to my attention (I am making no pretense here to an exhaustive survey of the literature) are: Blum and Kaufman (4) in 14 male patients with peptic ulcer; McNeil and Blum (7) relating handwriting signs with the psychosexual dimensions; and Ellis (6), who, in a patient with some two hundred hours of psychoanalysis, had the test rated by 22 other psychologists. In a survey published in 1952 by Blum and Hunt (3), they reported also an investigation in stutterers (Merchant); sexual offenders (Lindner); behavior in small groups (Swanson). A more recent investigation is that of Molish, Lyon, and Briggs (9), who compared Blacky test findings in 1947 naval recruits found "not maladjusted" by the Navy's neuropsychiatry service, with 390 recruits who were rejected as unsuitable. These investigators, while alert to the psychoanalytic implications of their results, were especially interested in the test's potency for differentiating their two groups. An inference from their research is that, whatever the agreements with psychoanalytic theory, the Blacky has its value in practical application, as derived from empirical data.

The test has thus a differentiating potency, and without a doubt it has the projective instrument value of opening a window to latent character traits. At the same time results are suspect, as I see it, by reason of a major fallacy in the technique of treating the data. This consists in the method of "scoring" each story and the other criteria as "strong" and the various gradations of "strong." The judg-

ments are made by each examiner, without recourse to any frame of reference. Scorings must, therefore, be a function of each scorer. They can, I will concede, be made reliable as between two investigators who have worked together closely and have learned the same signs and language, in recognizing and naming the psychological traits or "dimensions." Except between two such colleagues, or for some one examiner, results are only spuriously comparable. The degree of error to which the procedure is subject is seen in Blum's own monograph: a denial of a response is construed as a repression, and since a repression involves strong emotion, the response is scored "strong" (2, p. 34). To construe denial as repression—which is a hypothesis and not a datum—is interpretation. Thus a quantitative profile is developed and is reported as *S*'s production although it is actually the examiner's interpretation of *S*'s production.

This criticism touches on what is the Achilles' heel of each projective test. It dictates the need, therefore, of exploiting the objectivity potential in the test. The needed technique is that of patterning out its data into impersonal structure to be described operationally, and so creating some stable frames of reference by which the productions of any *S* can be judged, if not measured. Some of the projective tests do lend themselves more easily to such impersonal structuralization than others. But to the extent that the instrument falls short of such objectivity, it is not a test but a two-person situation, two persons—*S* and examiner, each of whom may influence the results. Schafer (10) has been especially emphatic concerning the interpersonal relation between examiner and subject, and the possible role of the examiner's personality in fashioning the final test picture. Brunswik writing from another theoretical viewpoint, arrives at a similar conclusion. "Since in any testing procedure the examiner constitutes part of the external stimulus situation, representative design demands that examiners should also be sampled," and he refers to the "double jeopardy of generalization" (5, p. 131).

An important implication follows from these positions of Schafer and of Brunswik: projective test investigations will need to find some method of keeping the examiner factor constant, if reports by two or more persons are to be comparable. The problem becomes at this point the baby of the theorists and logicians of science. This circumstance does not, however, ease the burden of the psychologists searching personality by means of projective tests. The fact is they have so far been arriving at right answers, even if they do not know how. Their answers are bound to be that much more accurate as they are based in sound method, and as they take pains to devise such. This is where I see Blacky and his growing pains at the present. The test has the promise warranting its developing. In spite of the fallacy in the scoring procedure, its agreement with clinical theory is logically sound. It can serve, further, as an example of other *ad hoc* techniques which can be devised to explore psychoanalytic theory.

References

1. Aronson, M. L. A study of the Freudian theory of paranoia by means of the Blacky pictures. *J. proj. Tech.*, 1953, **17**, 3–19.
2. Blum, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.*, 1949, **39**, 3–99.
3. Blum, G. S., & Hunt, H. F. The validity of the Blacky pictures. *Psychol. Bull.*, 1952, **49**, 238–250.
4. Blum, G. S., & Kaufman, Jewel B. Two patterns of personality dynamics in male peptic ulcer patients as suggested by responses to the Blacky pictures. *J. clin. Psychol.*, 1952, **8**, 273–278.
5. Brunswik, E. *Perception and representative design of psychological experiments.* Berkeley, Calif.: Univer. of California Press, 1956.
6. Ellis, A. The Blacky test used with a psychoanalytic patient. *J. clin. Psychol.*, 1953, **9**, 167–172.
7. McNeil, E., & Blum, G. S. Handwriting and psychosexual dimensions of personality. *J. proj. Tech.*, 1952, **16**, 476–484.
8. Michal-Smith, H., Hammer, E., & Spitz, H. The use of the Blacky pictures with a child whose Oedipal desires are close to consciousness. *J. clin. Psychol.*, 1951, **7**, 280–282.
9. Molish, H. B., Lyon, B., & Briggs, D. L. Character structure of adjusted and maladjusted naval recruits as measured by the Blacky pictures. *Amer. J. Orthopsychiat.*, 1954, **24**, 164–174.
10. Schafer, R. *Psychoanalytic interpretation in Rorschach testing.* New York: Grune & Stratton, 1954.
11. Seward, J. P. Psychoanalysis, deductive method, and the Blacky test. *J. abnorm. soc. Psychol.*, 1950, **45**, 529–535.

JOURNAL OF CONSULTING PSYCHOLOGY

Edited By
Laurance F. Shaffer
Teachers College, Columbia University

Advisory Editors

Donald W. Fiske
University of Chicago

Ann Magaret Garner
University of Illinois

Harrison G. Gough
University of California

William A. Hunt
Northwestern University

E. Lowell Kelly
University of Michigan

Morris Krugman
Board of Education
New York, N. Y.

Bertha M. Luckey
Board of Education
Cleveland, Ohio

Fred McKinney
University of Missouri

Anne Roe
New York, N. Y.

Carl R. Rogers
University of Chicago

R. Nevitt Sanford
University of California

Edward Joseph Shoben, Jr.
Teachers College
Columbia University

George J. Wischner
University of Pittsburgh

J. R. Wittenborn
Rutgers University

Managing Editor
Lorraine Bouthilet

VOLUME 20

1956

Published Bimonthly By

The American Psychological Association, Inc.
1333 Sixteenth St., N. W.
Washington 6, D. C.

Printed At
Lancaster, Pennsylvania

Copyright © 1956 by the American Psychological Association, Inc.

Contents of Volume 20

Altus, Grace T. A WISC Profile for Retarded Readers - - - - - - - - - - - - - - - 155
Altus, Grace T. Some Correlates of the Davis-Eells Tests - - - - - - - - - - - - - 227
Ansbacher, H. L. The "Alienation Syndrome" and Adler's Concept of "Distance" - - - 483
Barnes, E. H. Factors, Response Bias, and the MMPI - - - - - - - - - - - - - - 419
Barnes, E. H. Response Bias and the MMPI - - - - - - - - - - - - - - - - - - 371
Barratt, E. S. The Relationship of the Progressive Matrices (1938) and the Columbia Mental Maturity Scale to the WISC - - - - - - - - - - - - - - - - - - - 294
Beck, L. H. See *Rosvold, H. E.*
Bellugi, Ursula. See *Crandall, V. J.*
Bendig, A. W. Ranking Methodology: the Development of a Judgmental Criterion with Clinical Case Histories - 75
Bendig, A. W. Reliability of Case History Ratings and Intellectual Ability of Graduate Raters - 142
Bendig, A. W. The Development of a Short Form of the Manifest Anxiety Scale - - - - 384
Bendig, A. W. The Personality of Judges and their Agreement with Experts in Judging Clinical Case Histories - 422
Berg, J., and Polyot, C. J. The Influence of Color on Reactions to Incomplete Figures - - 9
Bernstein, L. The Examiner as an Inhibiting Factor in Clinical Testing - - - - - - - 287
Bieri, J., and Blacker, E. External and Internal Stimulus Factors in Rorschach Performance - 1
Blacker, E. See *Bieri, J.*
Blum, G. S. "Reliability of the Blacky Test": A Reply to Charen - - - - - - - - - 406
Bordin, E. S. See *Raush, H. L.*
Bosquet, K. T., and Stanley, W. C. Discriminative Powers of Rorschach Determinants in Children Referred to a Child Guidance Clinic - - - - - - - - - - - - - - - 17
Bransome, E. D., Jr. See *Rosvold, H. E.*
Breiger, B. The Use of the W-B Picture Arrangement Subtest as a Projective Technique - 132
Breiger, B. See *Lundin, W. H.*
Brody, Janice R. See *Levy, L. H.*
Broen, W. E., Jr. See *Wirt, R. D.*
Calvin, A. D., McGuigan, F. J., Tyrrell, Sybil, and Soyars, M. Manifest Anxiety and the Palmar Perspiration Index - 356
Carpenter, B. See *Wiener, M.*
Carpenter, Janeth T. See *Wiener, M.*
Carr, A. C. See *Goldstein, A. P.*
Cartwright, D. S. Note on "Changes in Psychoneurotic Patients with and without Psychotherapy" - 403
Cattell, R. B. Second-Order Personality Factors in the Questionnaire Realm - - - - - 411
Chambers, J. L. Perceptual Judgment and Associative Learning Ability of Schizophrenics and Nonpsychotics - 211
Charen, S. A Reply to Blum - 407
Charen, S. Reliability of the Blacky Test - - - - - - - - - - - - - - - - - - 16
Conger, J. J., Sawrey, W. L., and Krause, L. F. A Reanalysis of Beck's "Six Schizophrenias" - 83
Conger, J. J., Sawrey, W. L., and Krause, L. F. Reply to Dr. Stephenson - - - - - - 90
Copple, G. E. Effective Intelligence as Measured by an Unstructured Sentence-Completion Technique - 357
Crandall, V. J., and Bellugi, Ursula. Rating Personal Adjustment Through an Analysis of Social Reinforcements - 49

Cuddy, J. M. See *Levy, N. M.*
Dahlstrom, W. G. See *Goodstein, L. D.*
Dana, R. H. Cross Validation of Objective TAT Scoring - - - - - - - - - - - - - 33
Davids, A. Some Comments on Ansbacher's Note on the "Alienation Syndrome" and Adler's Concept of "Distance" - 485
Davids, A. The Influence of Ego-Involvement on Relations Between Authoritarianism and Intolerance of Ambiguity - 179
Deabler, H. L. See *Saucer, R. T.*
Diamond, M. D. The Ability of Schizophrenics to Modify Responses in an Interpersonal Situation - 441
Dibner, A. S. Cue-Counting: A Measure of Anxiety in Interviews - - - - - - - - - 475
Dittmann, A. T. See *Raush, H. L.*
Doppelt, J. E. Estimating the Full Scale Score on the Wechsler Adult Intelligence Scale from Scores on Four Subtests - 63
Downing, W. L. See *Lotsof, E. J.*
Dugan, R. D. See *Levy, L. H.*
Ellis, A. The Effectiveness of Psychotherapy with Individuals Who Have Severe Homosexual Problems - 191
Epstein, S., and Smith, R. Repression and Insight as Related to Reaction to Cartoons - - 391
Fairweather, G. W., Moran, L. J., and Morton, R. B. Efficiency of Attitudes, Fantasies, and Life History Data in Predicting Observed Behavior - - - - - - - - - - - 58
Fairweather, G. W. See *Moran, L. J.*
Fisher, S. Plausibility and Depth of Interpretation - - - - - - - - - - - - - - - 249
Fisher, S. See *Moran, L. J.*
Fordyce, W. E. Social Desirability in the MMPI - - - - - - - - - - - - - - - - 171
Frank, G. H. The Wechsler-Bellevue and Psychiatric Diagnosis: A Factor Analytic Approach - 67
Frederickson, W. K. See *Winter, W. D.*
Fricke, B. G. Response Set as a Suppressor Variable in the OIAS and MMPI - - - - - 161
Gaston, C. O. See *Taulbee, E. S.*
Gehman, I. H., and Matyas, R. P. Stability of the WISC and Binet Tests - - - - - - 150
Gieseking, C. See *Williams, H. L.*
Goldstein, A. P., and Carr, A. C. The Attitudes of the Mothers of Male Catatonic and Paranoid Schizophrenics Toward Child Behavior - - - - - - - - - - - - - - 190
Goodstein, L. D., and Dahlstrom, W. G. MMPI Difference Between Parents of Stuttering and Nonstuttering Children - 365
Gorham, D. R. Use of the Proverbs Test for Differentiating Schizophrenics from Normals 435
Grizwok, R., and Scodel, A. Some Psychological Correlates of Humor Preferences - - - - 42
Guze, S. B. See *Matarazzo, J. D.*
Harway, N. I. See *Raush, H. L.*
Hays, W. L. See *Raush, H. L.*
Heathers, Louise B. See *Merrill, R. M.*
Helfand, I. Role Taking in Schizophrenia - - - - - - - - - - - - - - - - - - - 37
Henry, Edith M., and Rotter, J. B. Situational Influences on Rorschach Responses - - - 457
Hodges, A. Double Alternation: A Measure of Intelligence - - - - - - - - - - - - 59
Horan, E. M. Word Association Frequency Tables of Mentally Retarded Children - - - - 22
Jessor, R. Social Values and Psychotherapy - - - - - - - - - - - - - - - - - 264
Johnson, Laverne C. Body Cathexis as a Factor in Somatic Complaints - - - - - - - 145
Jones, H. G. The Evaluation of the Significance of Differences Between Scaled Scores on the WAIS: The Perpetuation of a Fallacy - - - - - - - - - - - - - - - - - 319
Kagan, J., and Mussen, P. H. Dependency Themes on the TAT and Group Conformity - 29

Kendall, Barbara S., Mills, W. B., and Thale, T. Comparison of Two Methods of Electroshock in their Effect on Cognitive Functions - - - - - - - - - - - - - - - - 423
Kenny, D. T. The Influence of Social Desirability on Discrepancy Measures Between Real Self and Ideal Self - 315
Knopf, I. J. Rorschach Summary Scores in Differential Diagnosis - - - - - - - - - 99
Krause, L. F. See *Conger, J. J.*
Lakin, M. Certain Formal Characteristics of Human Figure Drawings by Institutionalized Aged and by Normal Children - 471
Langston, R. D. See *Murphy, D. B.*
Lawton, M. P. Stimulus Structure as a Determinant of the Perceptual Response - - - 351
Leiding, W. C. See *Norman, R. D.*
Leventhal, H. The Effects of Perceptual Training on the Rorschach *W* and *Z* Scores - - 93
Levin, M. L. Validation of the Lowenfeld Mosaic Test - - - - - - - - - - - - - - 239
Levine, M., and Meltzoff, J. Cognitive Inhibition and Rorschach Human Movement Responses - 119
Levy, L. H. A Note on Research Methodology Used in Testing for Examiner Influence in Clinical Test Performance - 286
Levy, L. H. The Meaning and Generality of Perceived Actual-Ideal Discrepancies - - - 396
Levy, L. H., Brody, Janice R., and Windman, Georgia O. The Relationship Between the Inferential Potential of Rorschach and TAT Protocols - - - - - - - - - - - - 27
Levy, L. H., and Dugan, R. D. A Factorial Study of Personal Constructs - - - - - - - 53
Levy, N. M., and Cuddy, J. M. Concept Learning in the Educationally Retarded Child of Normal Intelligence - 445
Litwin, Dorothy. See *Meltzoff, J.*
Lorr, M., and Rubinstein, E. A. Personality Patterns of Neurotic Adults in Psychotherapy - 257
Lotsof, E. J., and Downing, W. L. Two Measures of Anxiety - - - - - - - - - - - - 170
Lubin, A. See *Williams, H. L.*
Lundin, W. H., and Breiger, B. A Comparison of Productivity and Fantasy on the Rorschach and Projective Movement Sequences - - - - - - - - - - - - - - - - - - 342
McCraven, Vivian G. See *Singer, J. L.*
McGuigan, F. J. See *Calvin, A. D.*
Mann, L. The Relation of Rorschach Indices of Extratension and Introversion to a Measure of Responsiveness to the Immediate Environment - - - - - - - - - - - - - - 114
Maradie, L. J. The Goal-Spurt Hypothesis and the Rorschach Test - - - - - - - - 205
Mason, C. F. Preillness Intelligence of Mental Hospital Patients - - - - - - - - - 297
Matarazzo, J. D., Saslow, G., and Guze, S. B. Stability of Interaction Patterns during Interviews: A Replication - 267
Matarazzo, J. D. See *Matarazzo, Ruth G.*
Matarazzo, Ruth G., and Matarazzo, J. D. Anxiety Level and Pursuitmeter Performance - 70
Matyas, R. P. See *Gehman, I. H.*
Mehlman, B., and Vatovec, E. A Validation Study of the Bender-Gestalt - - - - - - - 71
Meltzoff, J., and Litwin, Dorothy. Affective Control and Rorschach Human Movement Responses - 463
Meltzoff, J. See *Levine, M.*
Merrill, R. M., and Heathers, Louise B. The Relation of the MMPI to the Edwards Personality Preference Schedule on a College Counseling Center Sample - - - - - - - - 310
Mills, W. B. See *Kendall, Barbara S.*
Mirsky, A. F. See *Rosvold, H. E.*
Moran, L. J., Fairweather, G. W., Fisher, S., and Morton, R. B. Psychological Concomitants to Rate of Recovery from Tuberculosis - - - - - - - - - - - - - - - - - 199

Moran, L. J., Fairweather, G. W., and Morton, R. B. Some Determinants of Successful and Unsuccessful Adaptation to Hospital Treatment of Tuberculosis - - - - - - - - 125
Moran, L. J. See *Fairweather, G. W.*
Morton, R. B. See *Fairweather, G. W.*
Morton, R. B. See *Moran, L. J.*
Murphy, D. B., and Langston, R. D. A Short Form of the Wechsler-Bellevue and the Army Classification Battery as Measures of Intelligence - - - - - - - - - - - - - - - 405
Murphy, D. G. The Revalidation of Diagnostic Tests for Alcohol Addiction - - - - - - 301
Mussen, P. H. See *Kagan, J.*
Nitsche, C. J., Robinson, J. F., and Parsons, E. T. Homosexuality and the Rorschach - - 196
Norman, R. D., and Leiding, W. C. The Relationship Between Measures of Individual and Mass Empathy - 79
Parsons, E. T. See *Nitsche, C. J.*
Pascal, G. R., and Zax, M. Psychotherapeutics: Success or Failure - - - - - - - - - 325
Peek, R. M., and Storms, L. H. Validity of the Marsh-Hilliard-Liechti MMPI Sexual Deviation Scale in a State Hospital Population - - - - - - - - - - - - - - - 133
Peters, G. A., Jr. Emotional and Intellectual Concomitants of Advanced Chronic Alcoholism - 390
Phares, E. J., and Rotter, J. B. An Effect of the Situation of Psychological Testing - - 291
Polyot, C. J. See *Berg, J.*
Psychological Test Reviews - - - - - - - - - - - - - - - - 157, 237, 321, 408, 487
Purcell, K. A Note on Porteus Maze and Wechsler-Bellevue Scores as Related to Antisocial Behavior - 361
Purcell, K. The TAT and Antisocial Behavior - - - - - - - - - - - - - - - - - - - 449
Raush, H. L., Sperber, Z., Rigler, D., Williams, J., Harway, N. I., Bordin, E. S., Dittman, A. T., and Hays, W. L. A Dimensional Analysis of Depth Interpretation - - - - - - 43
Reznikoff, M., and Tomblen, D. The Use of Human Figure Drawings in the Diagnosis of Organic Pathology - 467
Rigler, D. See *Raush, H. L.*
Roberts, Ruth. See *Sinnett, E. R.*
Robinson, H. A. See *Silverstein, A. B.*
Robinson, J. F. See *Nitsche, C. J.*
Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., Jr., and Beck, L. H. A Continuous Performance Test of Brain Damage - - - - - - - - - - - - - - - - - 343
Rotter, J. B. See *Henry, Edith M.*
Rotter, J. B. See *Phares, E. J.*
Rubinstein, E. A. See *Lorr, M.*
Rubinstein, I. See *Williams, H. L.*
Sarason, I. G. The Relationship of Anxiety and "Lack of Defensiveness" to Intellectual Performance - 220
Sarason, I. See *Rosvold, H. E.*
Saslow, G. See *Matarazzo, J. D.*
Saucer, R. T., and Deabler, H. L. Perception of Apparent Motion in Organics and Schizophrenics - 385
Sawrey, W. L. See *Conger, J. J.*
Scherer, I. W. See *Winne, J. F.*
Scodel, A. See *Griswok, R.*
Shontz, F. C. Evaluative Conceptualizations as the Basis for Cllnical Judgments - - - 223
Siegman, A. W. Cognitive, Affective, and Psychopathological Correlates of the Taylor Manifest Anxiety Scale - 137
Siegman, A. W. The Effect of Manifest Anxiety on a Concept Formation Task, a Nondirected Learning Task, and on Timed and Untimed Intelligence Tests - - - - - - 176

Silverstein, A. B., and Robinson, H. A. The Representation of Orthopedic Disability in Children's Figure Drawings - 333
Singer, J. L., Wilensky, H., and McCraven, Vivian G. Delaying Capacity, Fantasy, and Planning Ability: A Factorial Study of Some Basic Ego Functions - - - - - - - 375
Sinnett, E. R., and Roberts, Ruth. Rorschach Approach Type and the Organization of Cognitive Material - 109
Sipprelle, C. N., and Swensen, C. H. Relationship of Sexual Adjustment to Certain Characteristics of Human Figure Drawings - - - - - - - - - - - - - - - - - - 197
Sisson, B. D. See *Taulbee, E. S.*
Smith, R. See *Epstein, S.*
Soyars, M. See *Calvin, A. D.*
Sperber, Z. See *Raush, H. L.*
Stanley, W. C. See *Bosquet, K. T.*
Starer, E. The Use of the Kaleidoscope as an Adjunct to the Rorschach - - - - - - - 466
Stephenson, W. A Reply to the Reanalysis of Beck's "Six Schizophrenias" - - - - - - - 88
Storms, L. H. See *Peek, R. M.*
Swensen, C. H. See *Sipprelle, C. N.*
Taft, R. Intolerance of Ambiguity and Ethnocentrism - - - - - - - - - - - - - - - 153
Taulbee, E. S., Sisson, B. D., and Gaston, C. O. Affective Ratio and 8–9–10 Per Cent on the Rorschach Test for Normals and Psychiatric Groups - - - - - - - - - - - - - 105
Taylor, J. W. Relationship of Success and Length in Psychotherapy - - - - - - - - - 332
Thale, T. See *Kendall, Barbara S.*
Tolor, A. A Comparison of the Bender-Gestalt and the Digit-Span Test as Measures of Recall - 305
Tomblen, D. See *Reznikoff, M.*
Trites, D. K. Evaluation of Assumptions Underlying Interpretation of Sentence Completion Tests - 8
Tucker, J. E. Group Psychotherapy with Chronic Psychotic Soiling Patients - - - - - 430
Tyrrell, Sybil. See *Calvin, Allen D.*
Vatovec, E. See *Mehlman, B.*
Webb, W. B. A Procedure for Obtaining Self-Ratings and Group Ratings - - - - - - - 233
Wetherhorn, M. Flexor-Extensor Movement on the Rorschach - - - - - - - - - - - 204
Wickes, T. A., Jr. Examiner Influence in a Testing Situation - - - - - - - - - - - - 23
Wiener, M., Carpenter, B., and Carpenter, Janeth T. Determination of Defense Mechanisms for Conflict Areas from Verbal Material - - - - - - - - - - - - - - - - - - 215
Wilensky, H. See *Singer, J. L.*
Williams, H. L., Lubin, A., Gieseking, C., and Rubinstein, I. The Relation of Brain Injury and Visual Perception to Block Design Rotation - - - - - - - - - - - - - - - 275
Williams, J. See *Raush, H. L.*
Windman, Georgia O. See *Levy, L. H.*
Winne, J. F., and Scherer, I. W. A Second Study of Psychological Changes during the First Year Following Prefrontal Lobotomy - - - - - - - - - - - - - - - - - - 281
Winter, W. D., and Frederickson, W. K. The Short-Term Effects of Chlorpromazine on Psychiatric Patients - 431
Wirt, R. D. Actuarial Prediction - 123
Wirt, R. D. Ideational Expression of Hostile Impulses - - - - - - - - - - - - - - 185
Wirt, R. D., and Broen, W. E., Jr. The Relation of the Children's Manifest Anxiety Scale to the Concept of Anxiety as Used in the Clinic - - - - - - - - - - - - - - - - 482
Zax, M. See *Pascal, G. R.*
Zirkle, G. A. Learning in the Immediate Post-Electroshock Period - - - - - - - - - 399
Zweibelson, I. Test Anxiety and Intelligence Test Performance - - - - - - - - - - 479

Outstanding McGRAW-HILL *Books . . . .*

CLINICAL PSYCHOLOGY: The Study of Persons

By Richard W. Wallen, Creelman Associates. *McGraw-Hill Series in Psychology.* 400 pages, $6.00

This outstanding new book presents the major methods used by clinical psychologists at a level suitable for students who are beginning the study of the subject. The treatment of the interview, projective methods, psychometric tests, and psychotherapy not only gives a survey of the clinician's work, but establishes a foundation for further, specialized courses. Stress is laid on using many different kinds of information about persons in a creative way, rather than relying upon mechanical, standard interpretations. Case material helps students become more perceptive and think like professional clinicians.

BASIC CONCEPTS IN VOCATIONAL GUIDANCE

By Herbert Sanderson, University of Buffalo. 346 pages, $4.75

Covering the theoretical and practical aspects of counseling adolescents and adults, this comprehensive text aims to help students and practicing counselors familiarize themselves with the fundamental principles in educational and vocational counseling. It discusses the psychological forces that prompt the client to seek vocational help, the difficulties encountered in entering counseling work, the counseling process itself, the role of the counselor and his psychological needs, and the ending phase. Here is a text which concerns itself with such vital forces as the client's feelings, his reality, and his perception of the world of work.

PSYCHOANALYTIC THEORIES OF PERSONALITY

By Gerald S. Blum, University of Michigan. *McGraw-Hill Series in Psychology.* 219 pages, $4.25

The first integrated treatment of various psychoanalytic theories (propounded by Freud and other leading theorists) on personality. It constitutes an indispensable guide to the graduate student in clinical psychology and to the psychiatrist. The integration takes place by fitting the contribution of each theorist into a genetic sequence from before birth up to adulthood. The author evaluates the theories from a scientific research standpoint, including brief resumes of existing experimental data, suggestive evidence from related fields (cultural anthropology and learning theory), consideration of logical inconsistencies and semantic confusions, and comparisons of overlapping views.

PERSONALITY MEASUREMENT

By Leonard W. Ferguson, Research Associate, Life Insurance Agency Management Association. *McGraw-Hill Series in Psychology.* 445 pages, $6.50

Here is an account of the major methods by which personality tests have been constructed. From the discussions, the student will acquire new facility and perception in interpreting the scores which result from tests constructed by the various methods considered. The text brings together in one place a comprehensive account of the test construction methodologies with which all serious students of psychology must be familiar. Treatment involves a sufficiently detailed discussion of each method so that the student can acquire a real working knowledge of the basic and fundamental principles involved, and clearly and specifically differentiates between alternative tests in the same area.

Send for copies on approval

McGRAW-HILL BOOK CO., INC.

330 West 42nd Street, New York 36, N. Y.

THE DEVEREUX FOUNDATION

Since 1940, nonprofit sponsor of

DEVEREUX SCHOOLS

Established 1912

THERAPEUTIC SCHOOLS

VOCATIONAL COMMUNITIES

REMEDIAL CAMPS

SCHOOLS

EARL D. BOND DIVISION — Pennsylvania
WALTER L. TRIADWAY DIVISION — California

A complete scholastic program, in residence, for boys and girls presenting emotional difficulties, who can benefit by a program of therapeutic education, supplemented by individual psychotherapy where indicated.

Separate facilities and campuses for different age groups, from pre-kindergarten through junior college.

COMMUNITIES

LEO KANNER DIVISION — Pennsylvania
RANCH DIVISION — California

"Life-experience" and vocational programs for children, adolescents, and young adults with impaired intellectual or CNS function.

Separate, self-contained campuses for homogeneous groups, in terms of age and social maturity.

CAMPS

PENNSYLVANIA — MAINE — CALIFORNIA

Each Devereux School or Community has its own individual camp. Remedial academic programs, full recreational facilities, permit therapy to continue through the long summer holiday.

Professional inquiries will be welcomed by JOHN M. BARCLAY, *Director of Development*
Devereux Foundation, Devon, Pennsylvania

Devereux Schools

UNDER THE DEVEREUX FOUNDATION

HELENA T. DEVEREUX, *Director*

Professional Associate Directors

ROBERT L. BRIGDEN, Ph.D. — EDWARD L. FRENCH, Ph.D.
MICHAEL B. DUNN, Ph.D. — ROBERT T. GRATTAN, M.D.
J. CLIFFORD SCOTT, M.D.

Santa Barbara, California — **Devon, Pennsylvania**